

A27441 nucleolin - Chinese hamster
 N; Alternate names: nonchromosomal protein C23; nucleolar protein C23
 C; Species: *Cricetulus griseus* (Chinese hamster)
 C; Date: 19-Nov-1988 #sequence revision 19-Nov-1988 #text_change 23-Jul-1999
 C; Accession: A27441; A24808; A24070
 R; Lapeyre, B.; Bourbon, H.; Amalric, F.
 Proc. Natl. Acad. Sci. U.S.A. 84, 1472-1476, 1987
 A; Title: Nucleolin, the major nucleolar protein of growing eukaryotic cells
 A; Reference number: A27441; MUID:87175501; PMID:3470736
 A; Accession: A27441

A;Molecule type: mRNA
A;Residues: 1-713 <LAP>
R;lapeyre, B.; Amalric, F.; Ghaffari, S.H.; Venkatarama Rao, S.V.; Dumber, T.S.; Olson, J. Biol. Chem. 261, 9167-9173, 1986
A;Title: Protein and cDNA sequence of a glycine-rich, dimethylarginine-containing region
A;Reference number: A24808; MUID:86250857; PMID:3755137
A;Accession: A24808
A;Molecule type: mRNA
A;Residues: 522-541,'G','543-544','R','546-713 <LA2>
A;Cross-references: GB:M13757; NID:g191004; PIDN:AAA36961.1; PID:g387047
R;lapeyre, B.; Caizergues-Terrer, M.; Bouche, G.; Amalric, F.
Nucleic Acids Res. 13, 5805-5816, 1985
A;Title: Cloning of cDNA encoding a 100 kDa nucleolar protein (nucleoline) of Chinese hamster ovary cells
A;Reference number: A24070; MUID:85297777; PMID:2994013
A;Accession: A24070
A;Molecule type: mRNA
A;Residues: 179-238 <LA3>
C;Superfamily: nucleolin; ribonucleoprotein repeat homology
C;Keywords: DNA binding; nucleus; phosphoprotein
F;325-373/Domain: ribonucleoprotein repeat homology <RRM1>
F;394-456/Domain: ribonucleoprotein repeat homology <RRM2>
F;486-549/Domain: ribonucleoprotein repeat homology <RRM3>
F;572-636/Domain: ribonucleoprotein repeat homology <RRM4>

Query Match 38.7%; Score 79; DB 2; Length 713;
Best Local Similarity 45.7%; Pred.No. 5.8;
Matches 16; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 7 VSSKDKENISKENDVDLDEKEEAETEESELEEK 41
| : : : : | : : : : | : : : : | : : : : |
Db 237 VAEDDDDEEDEDDEEDEDDEEDEDDEEDEDDEE 271

RESULT 3
D85087
hypothetical protein ATg08710 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: D85087
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory Press
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: D85087
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-715 <STO>
C;Cross-references: GB:NC_001268; NID:g7267513; PIDN:CAB77996.1; GSFPDB:GN001440
C;Genetics:
A;Map position: 4

```

Query Match      38.7%; Score 79; DB 2; Length 715;
Best Local Similarity 42.9%; Pred.No 5.8;
Matches 15; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY      7 VSSKDKENIKSGNDVDLDEKEEAETEEEEELEEK 41
      ::::: ::::: ::::: ::::: :::::
DB      74 LTEQDPENVEEESSEEEKEEEKEEEEEESEEE 108

RESULT 4
D81798
lactoferrin-binding protein NMA1740 [imported] - Neisseria meningitidis (strain Z2491 se
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: D81798
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagele, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: D81798

```

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-741 <PAR>
A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB84968.1; PID:g7380381
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: lbpB, NMA1740

Query Match 38.7%; Score 79; DB 2; Length 741;
Best Local Similarity 51.5%; Pred. No. 6.1;
Matches 17; Conservative 8; Mismatches 6; Indels 2; Gaps 1;

Oy 8 SSKOKENISKENDVDLKEEAEETEEELER 40
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 510 NSEDEETIAEDDD--FAEEVEPEESPEE 540

RESULT 5
A35466
progesterone receptor form B - chicken
N:Contains: progesterone receptor form A
C:Species: Gallus gallus (Chicken)
C>Date: 18-Nov-1994 #sequence revision 18-Nov-1994 #text change 20-Aug-1999
C:Accession: A35466; S06284; A40903; A24661; A24312; A40911; A61552
R:Jeltsch, J.M.; Turcotte, B.; Garnier, J.M.; Lerouge, T.; Krozowski, Z.; Gronemeyer, H.;
J. Biol. Chem. 265, 3967-3974, 1990
A>Title: Characterization of multiple mRNAs originating from the chicken progesterone re
A:Reference number: A35466; MUID:90154085; PMID:2303488
A:Accession: A35466
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-786 <JEL>
A:Cross-references: GB:M32732; GB:J05240; NID:g212554; PIDN:AAA49011.1; PID:g212558
R:Gronemeyer, H.; Turcotte, B.; Quirin-Stricker, C.; Bocquel, M.T.; Meyer, M.E.; Krozowsk
EMBO J. 6, 3985-3994, 1987
A>Title: The chicken progesterone receptor: sequence, expression and functional analysis.
A:Reference number: S06284; MUID:88166640; PMID:3443098
A:Accession: S06284
A:Molecule type: DNA
A:Residues: 1-786 <GRO>
A:Cross-references: EMBL:Y00092; NID:g63744; PIDN:CAA68282.1; PID:g63745
R:Connely, O.M.; Dobson, A.D.W.; Tsai, M.J.; Beattie, W.G.; Toft, D.O.; Huckaby, C.S.; I
Mol. Endocrinol. 1, 517-525, 1987
A>Title: Sequence and expression of a functional chicken progesterone receptor.
A:Reference number: A40903; MUID:91042592; PMID:3153474
A:Accession: A40903
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-64,'E',65-786 <CON>
A:Cross-references: GB:M37518
R:Connely, O.M.; Sullivan, W.P.; Toft, D.O.; Birnbaumer, M.; Cook, R.G.; Maxwell, B.L.; I
Science 233, 767-770, 1986
A>Title: Molecular cloning of the chicken progesterone receptor.
A:Reference number: A24661; MUID:86289413; PMID:2426779
A:Accession: A24661
A:Molecule type: mRNA
A:Residues: 128-133,'E',135-147,'E',149-164 <CO2>
A>Note: the authors translated the codon CAG for residue 7 as Glu and CAG for residue 21
R:Jeltsch, J.M.; Krozowski, Z.; Quirin-Stricker, C.; Gronemeyer, H.; Simpson, R.J.; Gar
Proc. Natl. Acad. Sci. U.S.A. 83, 5424-5428, 1986
A>Title: Cloning of the chicken progesterone receptor.
A:Reference number: A24312; MUID:86287271; PMID:2426697
A:Accession: A24312
A:Molecule type: mRNA
A:Residues: 417-490 <JE2>
A:Cross-references: GB:M14280; NID:g212607; PIDN:AAA49039.1; PID:g212608
A>Note: amino acid and corresponding nucleotide sequences are also shown for three small
R:Birnbaumer, M.; Hinrichs-Rosello, M.V.; Cook, R.G.; Schrader, W.T.; O'Malley, B.W.
Mol. Endocrinol. 1, 249-259, 1987
A>Title: Chemical and antigenic properties of pure 108,000 molecular weight chick prosteg
A:Reference number: A40911; MUID:88288199; PMID:3453692
A:Accession: A40911
A>Status: preliminary

A:Molecule type: protein
 A:Residues: 128-133, 'E', 135-147, 'E', 149-164, 546-558 <BIR>
 R;Simpson, R.J.; Grego, B.; Govindan, M.V.; Gronemeyer, H.
 Mol. Cell. Endocrinol. 52, 177-184, 1987
 A:Title: Peptide sequencing of the chick oviduct progesterone receptor form B.
 A:Reference number: A61552; MUID:88005426; PMID:3635303
 A:Accession: A61552
 A:Molecule type: protein
 A:Residues: 136-153, 168-174, 195-228, 526-537, 'X', 539, 546-563 <SIM>
 C;Genetics:
 A:Introns: 400/2; 451/1; 490/1; 591/1; 639/2; 683/1; 735/3
 C:Superfamily: progesterone receptor; erba transforming protein homology
 F;Keywords: DNA binding; nucleus; steroid hormone receptor; zinc finger
 F;1-786/Product: progesterone receptor form B #status predicted <MA1>
 F;128-786/Product: progesterone receptor form A #status predicted <MA2>
 F;419-682/Domain: erba transforming protein homology <ERBA>
 F;421-441/Region: zinc finger
 F;457-481/Region: zinc finger

Query Match 38.7%; Score 79; DB 2; Length 786;
 Best Local Similarity 45.9%; Pred. No. 6.4;
 Matches 17; Conservative 9; Mismatches 9; Indels 2; Gaps 1;
 Qy 5 LYVSSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 42 LYPRSDDEEEENEB--EEEEEPOQREEEEEEE 76
 C:Genetics:
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2924 <BAR>
 A:Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1
 C:Genetics:
 A:Gene: EMP1
 A:Introns: 2476/3

RESULT 6
 T18378
 variant-specific surface protein 1 - malaria parasite (Plasmodium falciparum)
 N;Alternate names: erythrocyte membrane binding protein 1 (EMP1)
 C:Species: Plasmodium falciparum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C:Accession: T18378
 R;Baruch, D.I.; Pasloske, B.L.; Singh, H.B.; Bi, X.; Ma, X.C.; Feldman, M.; Taraschi, T.
 Cell 82, 77-87, 1995
 A:Title: Cloning the P. falciparum gene encoding PfEMP1, a malarial variant antigen and
 A:Reference number: Z18925; MUID:95330812; PMID:7541722
 A:Accession: T18378
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2924 <BAR>
 A:Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1
 C:Genetics:
 A:Gene: EMP1
 A:Introns: 2476/3

Query Match 38.7%; Score 79; DB 2; Length 2924;
 Best Local Similarity 45.5%; Pred. No. 23;
 Matches 15; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
 Qy 8 SSKDKENISKENDVLDKEKEAEETEELEEK 40
 Db 777 SSOEDDFEEEEEDEGEAEAEVQSEKTD 809
 C:Genetics:
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2924 <BAR>
 A:Cross-references: EMBL:U27338; NID:9914918; PID:9914919; PIDN:AAB60251.1
 C:Genetics:
 A:Gene: EMP1
 A:Introns: 2476/3

RESULT 7
 T49830
 hypothetical protein B24H17.160 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49830
 R;Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49830
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-231 <SCH>
 A:Cross-references: EMBL:ALJ56815; GSPDB:GN001116; NCSP:B24H17.160
 A:Experimental source: BAC clone B24H17; strain OR74A
 C:Genetics:

A:Gene: NCSP:B24H17.160
 A:Map position: 6
 A:Introns: 38/1; 102/3; 128/3

Query Match 38.2%; Score 78; DB 2; Length 231;
 Best Local Similarity 48.5%; Pred. No. 2.4;
 Matches 16; Conservative 8; Mismatches 9; Indels 0; Gaps 0;
 Qy 9 SKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 44 SKDEEEEEEDEEEEEEDEEEEEEDEEEEEE 76
 C:Genetics:
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-387 <BUL>
 C:Genetics:
 A:Gene: FlyBase:up
 A:Cross-references: FlyBase:FBgn0004169
 C:Keywords: muscle

Query Match 37.7%; Score 77; DB 2; Length 387;
 Best Local Similarity 45.7%; Pred. No. 4.9;
 Matches 15; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
 Qy 7 VSSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 347 VEDEDEDEDEDEDEDEDEDEDEDEDEDEE 381
 C:Genetics:
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-396 <FYR1>
 A:Cross-references: EMBL:X54504
 R;Fyrberg, E.A.
 submitted to the EMBL Data Library, August 1990
 A:Reference number: S29482
 A:Accession: S29482
 A:Molecule type: mRNA
 A:Residues: 1-191, 'A', 193-396 <FYR2>
 A:Cross-references: EMBL:X54504; NID:98739; PID:98740
 C:Genetics:
 A:Gene: FlyBase:up
 A:Cross-references: FlyBase:FBgn0004169
 A:Map position: X

Query Match 37.7%; Score 77; DB 2; Length 396;
 Best Local Similarity 45.7%; Pred. No. 5;
 Matches 16; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
 Qy 7 VSSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 347 VEDEDEDEDEDEDEDEDEDEDEDEDEDEE 381
 C:Genetics:
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-396 <FYR1>
 A:Cross-references: EMBL:X54504
 R;Fyrberg, E.A.
 submitted to the EMBL Data Library, August 1990
 A:Reference number: S29482
 A:Accession: S29482
 A:Molecule type: mRNA
 A:Residues: 1-191, 'A', 193-396 <FYR2>
 A:Cross-references: EMBL:X54504; NID:98739; PID:98740
 C:Genetics:
 A:Gene: FlyBase:up
 A:Cross-references: FlyBase:FBgn0004169
 A:Map position: X

RESULT 9
 S13251
 troponin T - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Sep-1998
 C:Accession: S13251; S29482
 R;Fyrberg, E.; Fyrberg, C.C.; Beall, C.; Saville, D.L.
 J. Mol. Biol. 216, 657-675, 1990
 A:Title: Drosophila melanogaster troponin-T mutations engender three distinct syndromes
 A:Reference number: S13251; MUID:91080155; PMID:2124273
 A:Accession: S13251
 A:Molecule type: mRNA
 A:Residues: 1-396 <FYR1>
 A:Cross-references: EMBL:X54504
 R;Fyrberg, E.A.
 submitted to the EMBL Data Library, August 1990
 A:Reference number: S29482
 A:Accession: S29482
 A:Molecule type: mRNA
 A:Residues: 1-191, 'A', 193-396 <FYR2>
 A:Cross-references: EMBL:X54504; NID:98739; PID:98740
 C:Genetics:
 A:Gene: FlyBase:up
 A:Cross-references: FlyBase:FBgn0004169
 A:Map position: X

Query Match 37.7%; Score 77; DB 2; Length 396;
 Best Local Similarity 45.7%; Pred. No. 5;
 Matches 16; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
 Qy 7 VSSKDKENISKENDVLDKEKEAEETEELEEK 41
 Db 347 VEDEDEDEDEDEDEDEDEDEDEDEDEDEE 381
 C:Genetics:
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-396 <FYR1>
 A:Cross-references: EMBL:X54504
 R;Fyrberg, E.A.
 submitted to the EMBL Data Library, August 1990
 A:Reference number: S29482
 A:Accession: S29482
 A:Molecule type: mRNA
 A:Residues: 1-191, 'A', 193-396 <FYR2>
 A:Cross-references: EMBL:X54504; NID:98739; PID:98740
 C:Genetics:
 A:Gene: FlyBase:up
 A:Cross-references: FlyBase:FBgn0004169
 A:Map position: X

[illegible]

RESULT 10

A54514
Glutamic acid-rich protein precursor - malaria parasite (Plasmodium falciparum)
N;Alternate names: GARP
C;Species: Plasmodium falciparum
C;Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 21-Jul-2000
C;Accession: A54514
R;Triglia, T.; Stahl, H.D.; Crewther, P.E.; Silva, A.; Anders, R.F.; Kemp, D.J.
Mol. Biochem. Parasitol. 33, 199-202, 1988
A;Title: Structure of a Plasmodium falciparum gene that encodes a glutamic acid-rich protein precursor
A;Reference number: A54514; MUID:89040048; PMID:2903445
A;Accession: A54514
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-678 <TRI>
X;Cross-references: GB:J03998; NID:gl60298; PIDN:AAA29605.1; PID:gl60299
C;Notes:
A;introns: 25/3
C;Superfamily: histone H1
C;Keywords: tandem repeat

Query Match 37.7%; Score 77; DB 2; Length 678;
Best Local Similarity 43.8%; Pred. No. 8.;
Matches 14; Conservative 10; Mismatches 8; Indels 0; Gaps 0;

Qy 10 KDKENISKENDVDLDEKEEEAEETEEEEELEK 41
: : : : : : : : : : : : : : : : : :

D**b**
D**b**

RESULT 11

C;Accession: S29795; S13983
 C;date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C;Species: chloroplast *Oenothera picensis* subsp. *picensis* (evening primrose)
 C;Keywords: chloroplast
 C;Genome: chloroplast
 C;References: EMBL:X64616; NID:g14334; PID:g14335
 C;Molecule type: DNA
 A;Molecule type: DNA
 A;Accession: S29795
 A;Reference number: S29795; MUID:93169690; PMID:8435856
 A;Title: In-frame length mutations associated with short tandem repeats are located in u
 Curr. Genet. 23, 265-270, 1993
 R;Nimzyk, R.; Schoendorf, T.; Hachtel, W.
 Hypoetical protein 2280 - evening primrose (*Oenothera picensis* subsp. *picensis*) chloro

Query Match 37.7%; Score 77; DB 2; Length 721;
Best Local Similarity 51.6%; Pred. No. 8.9;
Matches 16; Conservative 6; Mismatches 9; Indels

```
Qy 11 DKENISKENDVLDKEEEAEETEEEELEEK 41
      :: || :: || :: || :: || :: || ::
Db 555 EEEEEPKEEDELHEEEEEEEEEEEEEDE 585
```

RESULT 12

S65169
hypothetical protein YPL158c - yeast (*Saccharomyces cerevisiae*)
N;Alternate names: hypothetical protein P2570
C;Species: *Saccharomyces cerevisiae*
C;Date: 10-Dec-1994 #sequence_revision 31-May-1996 #text_change 19-Apr-2002
C;Accession: S65169; S69441
R;Burnelle, B.; Coeter, F.; Goffeau, A.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S65154
A;Accession: S65169
A;Molecule type: DNA
A;Residues: 1-758 <PUR>

A;Cross-references: EMBL:Z73514; NID:gl370335; PID:e247047; PID:gl370336; MIPS:YPL1580
A;Experimental source: strain S288C (AB972)
R;Purnelle, B.; Combiez, S.; Coster, F.; Naveau, F.; Goffeau, A.
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies
ologue to the human phosphotyrosyl phosphatase activator PTPA and a homologue to the plant
A;Reference number: S69428
A;Accession: S69441
A;Molecule type: DNA
A;Residues: 1-758 <PUW>
A;Cross-references: EMBL:X96770; NID:gl403537; PID:e239043; PID:gl403551
C;Genetics:
A;Cross-references: SGD:S0006079
A;Map position: 16L

Query Match 37.7%; Score 77; DB 2; Length 758;
Best Local Similarity 46.9%; Pred. No. 9.4;
Matches 15; Conservative 8; Mismatches 9; Indels

QY 10 KDKENISKENDVDLDEKEEAAETEEEEELEEK 41
: | : ||| |::||: |::|:
Db 661 EDDEEEEEENDEEDEDEDEDEDEDEDEEEKR 692

RESULT 13

T42363
 Hypothetical protein 48 - ateline herpesvirus 3 (strain 73)
 C:Species: ateline herpesvirus 3
 A:Variety: strain 73
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
 C:Accession: T42963
 R:Albrecht, J.C.; Fleckenstein, B.
 submitted to the EMBL Data Library, August 1998
 A:Description: Primary structure of the herpesvirus ateles genome.
 A:Reference number: Z22274
 A:Accession: T42963
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-792 <ALB>
 A:Cross-references: EMBL:AF083424; PIDN:AAC95573.1
 A:Experimental source: strain 73

Query Match 37.7%; Score 77; DB 2; Length 792;
Best Local Similarity 50.0%; Pred. No. 9.8;
Matches 16; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 10 K K E N I S K E N D V L D E K E E A B E T E E E E L E E K 41
||:| |:| ||:| |:| ||:| |:| ||:|

Db 455 K D E E N K K G D E E D E E D E E D E E E E E E E E E 486

RESULT 14

T30189 myelin transcription factor 1 homolog - mouse
N:Alternate names: zinc finger protein Mytl1
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30189
J:Kim, J.G.; Armstrong, R.C.; Robinsky, A.; Agoston, D.V.; Wiese, C.; Nagle
J. Neurosci. Res. 50, 272-290, 1997
A:Title: Myelin transcription factor 1 (Mytl1) of the oligodendrocyte lineage
A:Reference number: Z20766; MUID:98038874; PMID:9373037
A:Accession: T30189
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1182 <KIM>

Query Match 37.7%; Score 77; DB 2; Length 1182;
Best Local Similarity 40.0%; Pred. No. 14;
Matches 18; Conservative 9; Mismatches 8; Indels 10; Gaps 1;

QY 7 VSSKDKENISKEN-----DDVLDEKEEAEETEELEEK 41
 Db 125 VDREDEBEIEEEDDDDDGDDVVEEEDDDDDDEEEEEE 169

RESULT 15

T46608
 zinc finger protein Png-1 - mouse
 C/Species: Mus musculus (house mouse)
 C/Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 18-Feb-2000
 C/Accession: T46608
 R/Weiner, J.; Chun, J.
 J. Comp. Neurol. 381, 130-142, 1997
 A/Title: Png-1, a nervous system-specific zinc finger gene, identifies regions containin
 A/Reference number: 223102; MUID: 97276971; PMID: 9130664
 A/Accession: T46608
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1188 <WEI>
 A/Cross-references: EMBL:U06338; NID:gl835754; PID:gl835755; PIDN:AACS3157.1
 A/Experimental source: strain BALB/c
 A/Genetics:
 A/Name: Png-1

Query Match 37.7%; Score 77; DB 2; Length 1188;
 Best Local Similarity 40.9%; Pred. No. 14;
 Matches 18; Conservative 9; Mismatches 7; Indels 10; Gaps 1;

QY 8 SSKDKENISKEN-----DDVLDEKEEAEETEELEEK 41
 Db 128 SREDEBEIEEEDDDDDGDDVVEEEDDDDDDEEEEEE 171

Search completed: September 22, 2004, 18:10:02
 Job time : 16.5868 secs



Result No.	Score	Query Match	Length	DB	ID	Description
1	84	41.2	1121	1	MYT1_HUMAN	Q01538 homo sapien
2	82	40.2	3135	1	S230_PLAFO	Q08372 plasmodium
3	81	39.7	120	1	U222_ARATH	Q81ef3 arabidopsis
4	79	38.7	713	1	NUCL_MESAU	P08199 mesocricetu
5	79	38.7	786	1	PRGR_CHICK	P07812 gallus gall
6	77.5	38.0	627	1	BALA_XENLA	Q8uivr5 xenopus lae
7	77	37.7	396	1	TRT_DROME	P19151 drosophila
8	77	37.7	678	1	GAPF_PLAFL	P13816 plasmodium
9	77	37.7	721	1	YCF2_OENPI	P31568 oenothera p
10	77	37.7	758	1	YP58_YEAST	Q99299 saccharomyc
11	76.5	37.5	1124	1	TCF8_HUMAN	P37275 homo sapien
12	76	37.3	2492	1	ATRX_HUMAN	P46100 homo sapien
13	76	37.3	2492	1	ATRX_PANTR	P46100 homo sapien
14	76	37.3	2492	1	ATRX_PONPY	Q7Ygm4 pango troglod
15	75.5	37.0	213	1	NEUM_CARAU	Q7Ygm3 pongo pygma
16	75.5	37.0	1257	1	RBL1_HUMAN	P17691 caraseius a
17	75	36.8	739	1	DAXX_MOUSE	P29374 homo sapien
18	75	36.8	1043	1	TCF8_MESAU	Q35613 mus musculu
19	75	36.8	1099	1	NRX1_HUMAN	Q60542 mesocricetu
20	75	36.8	1263	1	BAT8_MOUSE	Q60721 homo sapien
21	74	36.3	180	1	HMG1_CRIGR	Q92148 mus musculu
22	74	36.3	214	1	HMG1_MOUSE	P07156 cricetus
23	74	36.3	301	1	TRT2_CHICK	P07155 mus musculu
24	74	36.3	417	1	CRT1_BOVIN	P02642 gallus gall
25	74	36.3	421	1	CRT2_BOVIN	P52193 bos taurus
26	74	36.3	630	1	YCF2_OENVI	P42918 bos taurus
27	74	36.3	1210	1	BAT8_HUMAN	P31569 oenothera v
28	73.5	36.0	202	1	IPPD_BOVIN	Q96kq7 homo sapien
29	73.5	36.0	321	1	ABRA_PLAFL	P07516 bos taurus
30	73	35.8	110	1	THYA_MOUSE	P23746 plasmodium
31	73	35.8	110	1	THYA_HUMAN	P06454 homo sapien
32	73	35.8	111	1	THYA_RAT	P26350 mus musculu
33	73	35.8	137	1	IPPD_PIG	P06302 rattus norv


```

SEQUENCE OF 35-713 FROM N.A., AND SEQUENCE OF 1-34.
MEDLINE=871175501; PubMed=3470736;
Lapeyre B., Bourbon H., Amalric F.;
RT "Nucleolin, the major nucleolar protein of growing eukaryotic cells:
an unusual protein structure revealed by the nucleotide sequence.";
Proc. Natl. Acad. Sci. U.S.A. 84:1472-1476(1987).
[2]
FUNCTION.
MEDLINE=88312631; PubMed=3409881;
RX Erard M.S., Belenguer P., Caizergues-Ferrer M., Pantaloni A.,
RA Amalric F.;
RT "A major nucleolar protein, nucleolin, induces chromatin
decondensation by binding to histone H1.";
Eur. J. Biochem. 175:525-530(1988).
CC -1- FUNCTION: Nucleolin is the major nucleolar protein of growing
eukaryotic cells. It is found associated with intranucleolar
chromatin and preribosomal particles. It induces chromatin
decondensation by binding to histone H1. It is thought to play a
role in pre-rRNA transcription and ribosome assembly.
CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -1- SIMILARITY: Contains 4 RNA recognition motif (RRM) domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M15825; AAA36966.1; -.
DR PIR; A27441; A27441.
DR PDB; 1FJ7; 16-OCT-00.
DR PDB; 1FJC; 16-OCT-00.
DR PDB; 1FJE; 03-JAN-01.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; rrm; 4.
DR SMART; SM00360; RRM; 4.
DR PROSITE; PS50102; RRM; 4.
DR PROSITE; PS00030; RRM_RNP_1; 3.
KW Nuclear protein; Phosphorylation; Methylation; DNA-binding; Repeat;
KW RNA-binding; 3D-structure.
FT INIT MET 0 0
FT FT DOMAIN 141 169 ASP/GLU-RICH (ACIDIC) .
FT FT DOMAIN 188 213 ASP/GLU-RICH (ACIDIC) .
FT FT DOMAIN 239 271 ASP/GLU-RICH (ACIDIC) .
FT FT DOMAIN 307 383 RNA-BINDING (RRM) 1.
FT FT DOMAIN 393 466 RNA-BINDING (RRM) 2.
FT FT DOMAIN 485 559 RNA-BINDING (RRM) 3.
FT FT DOMAIN 571 646 RNA-BINDING (RRM) 4.
FT FT DOMAIN 648 701 ARG/GLY/PHE-RICH.
FT FT DOMAIN 56 133 8 X 8 AA TANDEM REPEATS OF X-T-P-X-K-K-
X-X.
FT REPEAT 56 63 1.
FT REPEAT 73 80 2.
FT REPEAT 81 88 3.
FT REPEAT 89 96 4.
FT REPEAT 97 102 5 (INCOMPLETE) .
FT REPEAT 103 110 6.
FT REPEAT 118 125 7.
FT REPEAT 126 133 8.
FT REPEAT 217 221 X-T-P-X-K-K-X-X MOTIF.
FT FT MOD_RES 143 143 PHOSPHORYLATION.
FT FT MOD_RES 156 156 PHOSPHORYLATION.
FT FT MOD_RES 187 187 PHOSPHORYLATION.
FT FT MOD_RES 655 655 METHYLATION (DI-) .
FT FT MOD_RES 659 659 METHYLATION (DI-) .
FT FT MOD_RES 665 665 METHYLATION (DI-) .
FT FT MOD_RES 669 669 METHYLATION (DI-) .
FT FT MOD_RES 673 673 METHYLATION (DI-) .
FT FT MOD_RES 679 679 METHYLATION (DI-) .
FT FT MOD_RES 681 681 METHYLATION (DI-) .
FT FT MOD_RES 687 687 METHYLATION (DI-) .

```



```
FT DOMAIN 1 420 MODULATING, PRO-RICH.
FT DNA_BIND 421 486 NUCLEAR RECEPTOR-TYPE.
FT C4-TYPE 421 441 C4-TYPE.
FT ZN_FING 457 481 STEROID-BINDING.
FT DOMAIN 487 786 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 48 80 Missing (in isoform B and isoform B').
FT VARSPLIC 1 127 /FTID=VSP_003707.
FT VARSPLIC 452 458 QHNYLCA -> TISYHCS (in isoform A' and isoform B').
FT VARSPLIC 459 786 /FTID=VSP_003708.
FT VARSPLIC 459 786 Missing (in isoform A' and isoform B').
FT CONFLICT 58 58 E -> DD (IN REF. 2).
FT CONFLICT 480 480 K -> N (IN REF. 2).
FT CONFLICT 489 489 G -> A (IN REF. 2).
FT CONFLICT 577 577 R -> T (IN REF. 2).
FT CONFLICT 642 642 M -> I (IN REF. 2).
FT CONFLICT 786 AA; 85743 MW; 6595595950BC45ED9 CRC64;
SQ SEQUENCE 786 AA; 85743 MW; 6595595950BC45ED9 CRC64;

Query Match 38.7%; Score 79; DB 1; Length 786;
Best Local Similarity 45.9%; Pred. No. 4.9;
Matches 17; Conservative 9; Mismatches 9; Indels 2; Gaps 1;

QY 5 LYVSSKDKENISKENDVDLDEKEEAEETEELEE 41
DB 42 LYPRSDDEEEENEE--EEEEEPPQREEEEEE 76

RESULT 6
ID BAIA_XENLA STANDARD; PRT; 627 AA.
AC Q8UVR5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-utilizing chromatin assembly and remodel factor 1 (xACF1) (Fragment).
DE (Fragment).
GN ACF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21668941; PubMed=11809820;
RA MacCallum D.E., Losada A., Kobayashi R., Hirano T.;
RA "ISWI remodeling complexes in Xenopus egg extracts: identification as major chromosomal components that are regulated by INCENP-aurora B."; Mol. Biol. Cell 13:25-39(2002).
CC -1- FUNCTION: May play a role in transcriptional regulation.
CC -1- SUBUNIT: Together with p18 and p20 proteins, it forms the xenopus version of CHRAC.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- PTM: Phosphorylated in mitosis.
CC -1- SIMILARITY: Belongs to the WAL family.
CC -1- SIMILARITY: Contains 1 bromodomain.
CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DB EMBL; AF412332; AAL60160.1; -
DR InterPro; IPR001487; Bromodomain.
DR InterPro; IPR001965; Znf PHD.
DR Pfam; PF00439; bromodomain; 1.
DR Pfam; PF00628; PHD; 1.
```

```
DR PRINTS; PR00503; BROMODOMAIN.
DR SMART; SM00297; BROMO; 1.
DR SMART; SM00249; PHD; 1.
DR PROSITE; PS00633; BROMODOMAIN 1; 1.
DR PROSITE; PS00014; BROMODOMAIN 2; 1.
DR PROSITE; PS01359; ZF PHD 1; 1.
DR PROSITE; PS00016; ZF PHD 2; 1.
KW Transcription regulation; Bromodomain; Zinc-finger; Coiled coil;
KW Nuclear protein; Phosphorylation.
FT NON_TER 1 1
FT ZN_FING 222 272 PHD-TYPE.
FT DOMAIN 518 588 BROMODOMAIN.
FT DOMAIN 281 327 COILED COIL (POTENTIAL).
SQ SEQUENCE 627 AA; 70208 MW; 42C17587E531D380 CRC64;

Query Match 38.0%; Score 77.5; DB 1; Length 627;
Best Local Similarity 45.9%; Pred. No. 5.3;
Matches 17; Conservative 5; Mismatches 10; Indels 5; Gaps 1;

QY 4 HLYVSSKDKENISKENDVDLDEKEEAEETEELEE 40
DB 282 HRYSDMSD-----EEEEELDKKEEEEEEEQELSE 313

RESULT 7
ID TRT_DROME STANDARD; PRT; 396 AA.
AC PI9351;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Troponin T, skeletal muscle (Upheled protein) (Intended thorax DE protein).
GN UP OR INT.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Asynchronous muscle;
RX MEDLINE=89141761; PubMed=28522559;
RA Bullard B., Leonard K., Larkins A., Butcher G., Karlik C.,
RA Fryberg E.A.;
RA "Troponin of asynchronous flight muscle."; J. Mol. Biol. 204:621-637(1988).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=91080155; PubMed=2124273;
RA Fryberg E.A., Fryberg C.C., Beall C., Saville D.L.;
RA "Drosophila melanogaster troponin-T mutations engender three distinct syndromes of myofibrillar abnormalities."; J. Mol. Biol. 216:657-675(1990).
CC -1- FUNCTION: Troponin T is the tropomyosin-binding subunit of troponin, the thin filament regulatory complex which confers calcium-sensitivity to striated muscle actomyosin ATPase activity.
CC Mutations in troponin T engender three distinct syndromes of myofibrillar abnormalities.
CC -1- SIMILARITY: Belongs to the troponin T family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DB EMBL; X54504; CAA38366.1; -
DR FIR; S13251; S13251.
DR FlyBase; FBgn0004169; up.
```

[illegible]

RESULT 12
RT ATRX_HUMAN STANDARD; PRT; 2492 AA.
RL ATRX_HUMAN
AC P46100; P51068; Q15886; Q9H021; Q9NTS3;
DT 01-NOV-1995 (Rel. 32, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-WAR-2004 (Rel. 43, Last annotation update)
DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
DE nuclear protein) (XNP) (Znf-HX).
GN ATRX OR RAD54L OR XH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4 AND 5), VARIANTS SER-596;
RP GLU-740 AND SER-1860, AND VARIANTS ATR-X.
RX MEDLINE=97123494; PubMed=8968741;
RA Picketts D.J., Higgs D.R., Bachoo S., Blake D.J., Quarrell O.W.J.,
RA Gibbons R.J.;
RT "ATRX encodes a novel member of the SNF2 family of proteins: mutations
RT point to a common mechanism underlying the ATR-X syndrome.";
RL Hum. Mol. Genet. 5:1899-1907(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 4).
RX MEDLINE=97386582; PubMed=9244431;
RA Villard L., Lossi A.-M., Cardoso C., Proud V., Chiaroni P.,
RA Collea L., Schwartz C., Fontes M.;
RT "Determination of the genomic structure of the XNP/ATRX gene encoding
RT a potential zinc finger helicase.";
RL Genomics 43:149-155(1997).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS SER-596 AND GLU-740.
RX MEDLINE=22763540; PubMed=1277533;
RA Kitano T., Schwarz C., Nickel B., Paabo S.;
RT "Gene diversity patterns at 10 X-chromosomal loci in humans and
RT chimpanzees.";
RL Mol. Biol. Evol. 20:1281-1289(2003).
RN [4]
RP SEQUENCE OF 860-2492 FROM N.A.
RX MEDLINE=95179111; PubMed=7874112;
RA Stayton C.L., Dabovic B., Guliasso M., Gecz J., Broccoli V.,
RA Giovanazzi S., Bosolasco M., Monaco L., Raetan S., Boncinelli E.,
RA Bianchi M.E., Consalez G.G.;
RT "Cloning and characterization of a new human Xq13 gene, encoding a
RT putative helicase.";
RL Hum. Mol. Genet. 3:1957-1964(1994).
RN [5]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=94214473; PubMed=8182050;
RA Gecz J., Pollard H., Consalez G., Villard L., Stayton C.L.,
RA Milasseau P., Khrestchatsky M., Fontes M.;
RT "Cloning and expression of the murine homologue of a putative human
RT X-linked nuclear protein gene closely linked to PGK1 in Xq13.3";
RL Hum. Mol. Genet. 3:39-44(1994).
RN [6]
RP SEQUENCE OF 2401-2492 FROM N.A., AND VARIANTS ATR-X.
RX MEDLINE=95211835; PubMed=7697714;
RA Gibbons R.J., Picketts D.J., Villard L., Higgs D.R.;
RT "Mutations in a putative global transcriptional regulator cause X-
RT linked mental retardation with alpha-thalassemia (ATR-X syndrome).";
RL Cell 80:837-845(1995).
RN [7]
RP SEQUENCE OF 1375-2492 FROM N.A.
RA Pearce A., Chapman J.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [8]
RP EZH2 BINDING.
RX MEDLINE=98167853; PubMed=9499421;
RA Cardoso C., Timsit S., Villard L., Khrestchatsky M., Fontes M.,
RA Collea L.;
RT "Specific interaction between the XNP/ATRX gene product and the SET

RT domain of the human EZH2 protein.";
RL Hum. Mol. Genet. 7:679-684(1998).
RN [9]
RP SUBCELLULAR LOCATION, AND ASSOCIATION WITH PERICENTROMERIC
RP HETEROCHROMATIN.
RX MEDLINE=20040663; PubMed=10570185;
RA McDowell T.L., Gibbons R.J., Sutherland H., O'Rourke D.M.,
RA Bickmore W.A., Pombo A., Turley H., Gatter K., Picketts D.J.,
RA Buckle V.J., Chapman L., Rhodes D., Higgs D.R.;
RT "Localization of a putative transcriptional regulator (ATRX) at
RT pericentromeric heterochromatin and the short arms of acrocentric
RT chromosomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13983-13988(1999).
RN [10]
RP DISEASE.
RX MEDLINE=20213147; PubMed=10751095;
RA Villard L., Fontes M., Ades L.C., Gecz J.;
RT "Identification of a mutation in the XNP/ATRX gene in a family
RT reported as Smith-Fineman-Myers syndrome.";
RL Am. J. Med. Genet. 91:83-85(2000).
RN [11]
RP DISEASE.
RX MEDLINE=22804448; PubMed=12858175;
RA Gibbons R.J., Pellagatti A., Garrick D., Wood W.G., Malik N.,
RA Ayub H., Langford C., Boulwood J., Wainscoat J.S., Higgs D.R.;
RT "Identification of acquired somatic mutations in the gene encoding
RT chromatin-remodeling factor ATRX in the alpha-thalassemia
RT myelodysplasia syndrome (ATMDS).";
RL Nat. Genet. 34:446-449(2003).
RN [12]
RP VARIANT ATR-X SER-1713.
RX MEDLINE=97196774; PubMed=9043863;
RA Villard L., Lacombe D., Fontes M.;
RT "A point mutation in the XNP gene, associated with an ATR-X phenotype
RT without alpha-thalassemia.";
RL Eur. J. Hum. Genet. 4:316-320(1996).
RN [13]
RP VARIANT JM GLN-2131.
RX MEDLINE=96224392; PubMed=8630485;
RA Villard L., Gecz J., Mattel J.-F., Fontes M., Saugier-Verber P.,
RA Munnich A., Lyonnet S.;
RT "XNP mutation in a large family with Juberg-Marsidi syndrome.";
RL Nat. Genet. 12:359-360(1996).
RN [14]
RP VARIANTS ATR-X.
RX MEDLINE=97467722; PubMed=9326931;
RA Gibbons R.J., Bachoo S., Picketts D.J., Aftimos S., Asenbauer B.,
RA Bergoffen J., Berry S.A., Dahl N., Fryer A., Keppler K., Kurosawa K.,
RA Levin M.L., Masuno M., Neri G., Pierpont M.E., Slaney S.F.,
RA Higgs D.R.;
RT "Mutations in transcriptional regulator ATRX establish the functional
RT significance of a PHD-like domain.";
RL Nat. Genet. 17:146-148(1997).
RN [15]
RP VARIANT ATR-X LEU-246.
RX MEDLINE=20123062; PubMed=10660327;
RA Fichera M., Romano C., Castiglia L., Failla P., Ruberto C., Amata S.,
RA Greco D., Cardoso C., Fontes M., Ragusa A.;
RT "New mutations in XNP/ATRX gene: a further contribution to
RT genotype/phenotype relationship in ATR/X syndrome.";
RL Hum. Mutat. 12:214-214(1998).
RN [16]
RP VARIANT SHS LYS-1742.
RX MEDLINE=99347960; PubMed=10417298;
RA Lossi A.-M., Millan J.M., Villard L., Orellana C., Cardoso C.,
RA Prieto F., Fontes M., Martinez F.;
RT "Mutation of the XNP/ATRX gene in a family with severe mental
RT retardation, spastic paraplegia and skewed pattern of X inactivation:
RT demonstration that the mutation is involved in the inactivation
RT bias.";
RL Am. J. Hum. Genet. 65:558-562(1999).
RN [17]
RP VARIANT CWS THR-2050.

RX MEDLINE=99326063; PubMed=10398237;
 RA Abidi F., Schwartz C.E., Carpenter N.J., Villard L., Fontes M.,
 RA Curtis M.;
 RT "Carpenter-Waziri syndrome results from a mutation in XNP.";
 RL Am. J. Med. Genet. 85:249-251(1999).
 RN [18]
 RP VARIANTS ATR-X GLU-175; 178-VAL--LYS-198 DEL; SER-190; PRO-219;
 RP LEU-246 AND CYS-249.
 RX MEDLINE=99219535; PubMed=10204841;
 RA Villard L., Bonino M.-C., Abidi F., Ragusa A., Belouigne J.,
 RA Loessi A.-M., Seaver L., Bonnefont J.-P., Romano C., Fichera M.,
 RA Lacombe D., Hanauer A., Philip N., Schwartz C.E., Fontes M.;
 RT "Evaluation of a mutation screening strategy for sporadic cases of
 RT ATR-X syndrome.";
 RL J. Med. Genet. 36:183-186(1999).
 RN [19]
 RP VARIANTS ATR-X SER-179; LEU-190; ILE-194; CYS-246; PHE-1552; SER-1645
 RP AND CYS-1847.
 RX MEDLINE=20451413; PubMed=10995512;
 RA Wada T., Kubota T., Fukushima Y., Saitoh S.;
 RA "Molecular genetic study of Japanese patients with X-linked alpha-
 RA thalassemia/mental retardation syndrome (ATR-X).";
 RL Am. J. Med. Genet. 94:242-248(2000).
 RN [20]
 RP VARIANT ATR-X MET-1621.
 RX MEDLINE=22108790; PubMed=12116232;
 RA Yntema H.G., Poppelars F.A., Derksen E., Oudakker A.R.,
 RA van Roosmalen T., Jacobs A., Obbema H., Brunner H.G., Hamel B.C.J.,
 RA "Expanding phenotype of XNP mutations: mild to moderate mental
 RT retardation.";
 RL Am. J. Med. Genet. 110:243-247(2002).
 CC -1- FUNCTION: Could be a global transcriptional regulator. Modifies
 CC gene expression by affecting chromatin. May be involved in brain
 CC development and facial morphogenesis.
 CC -1- SUBUNIT: Probably binds EZH2. Binds annexin V in a calcium and
 CC phosphatidylcholine/phosphatidylserine-dependent manner (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear. Associated with pericentromeric
 CC heterochromatin during interphase and mitosis, probably by
 CC interacting with HPI.
 CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
 CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB102642; BAC81111.1; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR000330; SNF2_N.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00176; SNF2_N; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 KW DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding;
 KW Zinc-finger.
 FT ZN_FING. 220 268
 FT NP_BIND 1594 1601
 FT SITE 1719 1722
 FT DOMAIN 745 750
 FT DOMAIN 1151 1156
 FT DOMAIN 1166 1169
 FT DOMAIN 1202 1206
 FT DOMAIN 1259 1266
 FT DOMAIN 1443 1466
 FT DOMAIN 1499 1502
 FT DOMAIN 1929 1939
 FT DOMAIN 1941 1948
 FT DOMAIN 2222 2225
 FT DOMAIN 2262 2265
 FT DOMAIN 2420 2425
 SQ SEQUENCE 2492 AA; 282567 MW; 56978AC7D37400DC CRC64;

 Query Match 37.3%; Score 76; DB 1; Length 2492;
 Best Local Similarity 52.9%; Pred. No. 27;
 Matches 18; Conservative 5; Mismatches 7; Indels 4; Gaps 1;

 QY 8 SSKDKENISKENDVDLDEKEEAETEELEEK 41
 Db 1435 SSENKSNSEEE-----EKEEEEEEEEEEEEE 1464

 RESULT 13
 ATRX_PANTR STANDARD; PRT; 2492 AA.
 AC QYQM4;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Transcriptional regulator ATRX (X-linked helicase II) (X-linked
 DE nuclear protein) (XNP).
 GN ATRX.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22763540; PubMed=12777533;
 RA Kitano T., Schwarz C., Nickel B., Paabo S.;
 RT "Gene diversity patterns at 10 X-chromosomal loci in humans and
 RT chimpanzees.";
 RL Mol. Biol. Evol. 20:1281-1289(2003).
 CC -1- FUNCTION: Could be a global transcriptional regulator. Modifies
 CC gene expression by affecting chromatin. May be involved in brain
 CC development and facial morphogenesis.
 CC -1- SUBUNIT: Probably binds EZH2. Binds annexin V in a calcium and
 CC phosphatidylcholine/phosphatidylserine-dependent manner (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Nuclear. Associated with pericentromeric
 CC heterochromatin during interphase and mitosis, probably by
 CC interacting with HPI.
 CC -1- SIMILARITY: Belongs to the SNF2/RAD54 helicase family.
 CC -1- SIMILARITY: Contains 1 PHD-type zinc finger.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB102642; BAC81111.1; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR000330; SNF2_N.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00176; SNF2_N; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 KW DNA repair; Nuclear protein; DNA-binding; Helicase; ATP-binding;
 KW Zinc-finger.
 FT ZN_FING. 220 268
 FT NP_BIND 1594 1601
 FT SITE 1719 1722
 FT DOMAIN 745 750
 FT DOMAIN 1151 1156
 FT DOMAIN 1166 1169
 FT DOMAIN 1202 1206
 FT DOMAIN 1259 1266
 FT DOMAIN 1443 1466
 FT DOMAIN 1499 1502
 FT DOMAIN 1929 1939
 FT DOMAIN 1941 1948
 FT DOMAIN 2222 2225
 FT DOMAIN 2262 2265
 FT DOMAIN 2420 2425
 SQ SEQUENCE 2492 AA; 282567 MW; 56978AC7D37400DC CRC64;

 Query Match 37.3%; Score 76; DB 1; Length 2492;
 Best Local Similarity 52.9%; Pred. No. 27;
 Matches 18; Conservative 5; Mismatches 7; Indels 4; Gaps 1;

 QY 8 SSKDKENISKENDVDLDEKEEAETEELEEK 41
 Db 1435 SSENKSNSEEE-----EKEEEEEEEEEEEEE 1464

 RESULT 14
 ATRX_PONPY



Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	41	14	US-10-294-770-14
2	204	100.0	41	16	US-10-774-602-1
3	101	49.5	28	14	US-10-294-770-1
4	101	49.5	28	16	US-10-774-602-1
5	97	47.5	28	14	US-10-294-770-4
6	97	47.5	28	16	US-10-238-741-4
7	97	47.5	28	16	US-10-774-602-4
8	97	47.5	64	14	US-10-294-770-1
9	97	47.5	64	16	US-10-238-741-1
10	97	47.5	64	16	US-10-774-602-1
11	87	42.6	1077	15	US-10-099-322-110
12	87	42.6	1077	15	US-10-044-564-110
13	83	40.7	197	13	US-10-101-487-51
14	83	40.7	197	13	US-10-101-487-114
15	80	39.2	161	16	US-10-437-963-108765

RESULT 2
US-10-774-602-14
; Sequence 14, Application US/10774602

Publication No. US20040141987A1
 GENERAL INFORMATION:
 APPLICANT: DRUILHE, PIERRE
 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 FILE REFERENCE: 248791USODIV
 CURRENT APPLICATION NUMBER: US/10/774,602
 CURRENT FILING DATE: 2004-02-10
 PRIOR APPLICATION NUMBER: US 09/356,947
 PRIOR FILING DATE: 1999-07-19
 PRIOR APPLICATION NUMBER: US 10/238,741
 PRIOR FILING DATE: 2002-09-11
 PRIOR APPLICATION NUMBER: US 08/416,711
 PRIOR FILING DATE: 1995-08-08
 PRIOR APPLICATION NUMBER: PCT/FR93/01024
 PRIOR FILING DATE: 1993-10-18
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 14
 LENGTH: 41
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic Peptide
 US-10-774-602-14

Query Match 100.0%; Score 204; DB 16; Length 41;
 Best Local Similarity 100.0%; Pred. No. 5.9e-13;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKENDVLDKEEAEETEESELEK 41
 DB 1 MLSHLYVSSKDKENISKENDVLDKEEAEETEESELEK 41

RESULT 3
 US-10-294-770-13
 Sequence 13, Application US/10294770
 Publication No. US20030161840A1
 GENERAL INFORMATION:
 APPLICANT: DRUILHE, PIERRE
 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 FILE REFERENCE: 230759USOCIP
 CURRENT APPLICATION NUMBER: US/10/294,770
 CURRENT FILING DATE: 2002-11-15
 PRIOR APPLICATION NUMBER: US 09/356,947
 PRIOR FILING DATE: 1999-07-19
 PRIOR APPLICATION NUMBER: US 08/416,711
 PRIOR FILING DATE: 1995-08-08
 PRIOR APPLICATION NUMBER: PCT/FR93/01024
 PRIOR FILING DATE: 1993-10-18
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 13
 LENGTH: 28
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic Peptide
 US-10-294-770-13

Query Match 49.5%; Score 101; DB 14; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKEND 20
 DB 9 MLSHLYVSSKDKENISKEND 28

RESULT 4
 US-10-774-602-13
 Sequence 13, Application US/10774602
 Publication No. US20040141987A1

GENERAL INFORMATION:
 APPLICANT: DRUILHE, PIERRE
 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 FILE REFERENCE: 248791USODIV
 CURRENT APPLICATION NUMBER: US/10/774,602
 CURRENT FILING DATE: 2004-02-10
 PRIOR APPLICATION NUMBER: US 09/356,947
 PRIOR FILING DATE: 1999-07-19
 PRIOR APPLICATION NUMBER: US 10/238,741
 PRIOR FILING DATE: 2002-09-11
 PRIOR APPLICATION NUMBER: US 08/416,711
 PRIOR FILING DATE: 1995-08-08
 PRIOR APPLICATION NUMBER: PCT/FR93/01024
 PRIOR FILING DATE: 1993-10-18
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 13
 LENGTH: 28
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Synthetic Peptide
 US-10-774-602-13

Query Match 49.5%; Score 101; DB 16; Length 28;
 Best Local Similarity 100.0%; Pred. No. 0.0032;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKEND 20
 DB 9 MLSHLYVSSKDKENISKEND 28

RESULT 5
 US-10-294-770-4
 Sequence 4, Application US/10294770
 Publication No. US20030161840A1
 GENERAL INFORMATION:
 APPLICANT: DRUILHE, PIERRE
 TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 FILE REFERENCE: 230759USOCIP
 CURRENT APPLICATION NUMBER: US/10/294,770
 CURRENT FILING DATE: 2002-11-15
 PRIOR APPLICATION NUMBER: US 09/356,947
 PRIOR FILING DATE: 1999-07-19
 PRIOR APPLICATION NUMBER: US 08/416,711
 PRIOR FILING DATE: 1995-08-08
 PRIOR APPLICATION NUMBER: PCT/FR93/01024
 PRIOR FILING DATE: 1993-10-18
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 28
 TYPE: PRT
 ORGANISM: Plasmodium falciparum
 US-10-294-770-4

Query Match 47.5%; Score 97; DB 14; Length 28;
 Best Local Similarity 95.0%; Pred. No. 0.0078;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLSHLYVSSKDKENISKEND 20
 DB 9 MLSHLYVSSKDKENISKENE 28

RESULT 6
 US-10-238-741-4
 Sequence 4, Application US/10238741
 Publication No. US20040096466A1
 GENERAL INFORMATION:
 APPLICANT: DRUILHE, PIERRE
 BOUHARCUN-TAYOUN, HASNAQ

OEUVRAY, CLAUDE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/238,741
FILING DATE: 09-Nov-2002
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/356,497
FILING DATE: 19-Jul-1999
APPLICATION NUMBER: US/08/416,711
FILING DATE: 08-AUG-1995
APPLICATION NUMBER: PCT/FR93/01024
FILING DATE: 18-OCT-1993
APPLICATION NUMBER: FR 92/12488
FILING DATE: 19-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-085-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-238-741-4
Query Match 47.5%; Score 97; DB 16; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.0078;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28
RESULT 7
US-10-774-602-4
Sequence 4, Application US/10774602
Publication No. US20040141987A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
FILE REFERENCE: 248791USODIV
CURRENT APPLICATION NUMBER: US/10/774,602
CURRENT FILING DATE: 2004-02-10
PRIOR APPLICATION NUMBER: US 09/356,947
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: US 10/238,741
PRIOR FILING DATE: 2002-09-11
PRIOR APPLICATION NUMBER: US 08/416,711
PRIOR FILING DATE: 1995-08-08
PRIOR APPLICATION NUMBER: PCT/FR93/01024

PRIOR FILING DATE: 1993-10-18
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent in version 3.1
SEQ ID NO 4
LENGTH: 28
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-774-602-4
Query Match 47.5%; Score 97; DB 16; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.0078;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28
RESULT 8
US-10-294-770-1
Sequence 1, Application US/10294770
Publication No. US20030161840A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
FILE REFERENCE: 230759USOCIP
CURRENT APPLICATION NUMBER: US/10/294,770
CURRENT FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US 09/356,947
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: US 08/416,711
PRIOR FILING DATE: 1995-08-08
PRIOR APPLICATION NUMBER: PCT/FR93/01024
PRIOR FILING DATE: 1993-10-18
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patent in version 3.1
SEQ ID NO 1
LENGTH: 64
TYPE: PRT
ORGANISM: Plasmodium falciparum
US-10-294-770-1
Query Match 47.5%; Score 97; DB 14; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.018;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MLSHLYVSSKDKENISKEND 20
Db 45 MLSHLYVSSKDKENISKENE 64
RESULT 9
US-10-238-741-1
Sequence 1, Application US/10238741
Publication No. US20040096466A1
GENERAL INFORMATION:
APPLICANT: DRUILHE, PIERRE
BOUHAROUN-TAYOUN, HASNAQ
OEUVRAY, CLAUDE
TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/10/238,741
 FILING DATE: 09-Nov-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/356,497
 FILING DATE: 19-Jul-1999
 APPLICATION NUMBER: US/08/416,711
 FILING DATE: 08-Aug-1995
 APPLICATION NUMBER: PCT/FR93/01024
 FILING DATE: 18-Oct-1993
 APPLICATION NUMBER: FR 92/12488
 FILING DATE: 19-Oct-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: OBLON, NORMAN F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 660-085-0 PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-413-3000
 TELEFAX: 703-413-2220
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 64 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-238-741-1

Query Match 47.5%; Score 97; DB 16; Length 64;
 Best Local Similarity 95.0%; Pred. No. 0.018;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
 DB 45 MLSHLYVSSKDKENISKENE 64

RESULT 10
 US-10-774-602-1
 ; Sequence 1, Application US/10774602
 ; Publication No. US20040141987A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUILHE, PIERRE
 ; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 ; FILE REFERENCE: 248791USODIV
 ; CURRENT APPLICATION NUMBER: US/10/774,602
 ; CURRENT FILING DATE: 2004-02-10
 ; PRIOR APPLICATION NUMBER: US 09/356,947
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: US 10/238,741
 ; PRIOR FILING DATE: 2002-09-11
 ; PRIOR APPLICATION NUMBER: US 08/416,711
 ; PRIOR FILING DATE: 1995-08-08
 ; PRIOR APPLICATION NUMBER: PCT/FR93/01024
 ; PRIOR FILING DATE: 1993-10-18
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 64
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 US-10-774-602-1

Query Match 47.5%; Score 97; DB 16; Length 64;
 Best Local Similarity 95.0%; Pred. No. 0.018;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
 DB 45 MLSHLYVSSKDKENISKENE 64

DB 45 MLSHLYVSSKDKENISKENE 64
 RESULT 11
 US-10-099-322-110
 ; Sequence 110, Application US/10099322
 ; Publication No. US20030215449A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes et al.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-240CIP
 ; CURRENT APPLICATION NUMBER: US/10/099,322
 ; CURRENT FILING DATE: 2002-09-11
 ; PRIOR APPLICATION NUMBER: 60/261,014
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,018
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/318,410
 ; PRIOR FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/261,013
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,026
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,029
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/313,170
 ; PRIOR FILING DATE: 2001-08-17
 ; PRIOR APPLICATION NUMBER: 10/044,564
 ; PRIOR FILING DATE: 2002-01-11
 ; NUMBER OF SEQ ID NOS: 324
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 110
 ; LENGTH: 1077
 ; TYPE: PRT
 ; ORGANISM: Drosophila melanogaster
 US-10-099-322-110

Query Match 42.6%; Score 87; DB 15; Length 1077;
 Best Local Similarity 55.9%; Pred. No. 3.1;
 Matches 19; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 8 SSKDKENISKENDVDLDEKEEAETEEEEK 41
 DB 970 ASKSDDESESENDDEDEEESESESEEEK 1003

RESULT 12
 US-10-044-564-110
 ; Sequence 110, Application US/10044564
 ; Publication No. US20040018196A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mezes et al.
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-240
 ; CURRENT APPLICATION NUMBER: US/10/044,564
 ; CURRENT FILING DATE: 2002-09-09
 ; PRIOR APPLICATION NUMBER: 60/261,014
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,018
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/318,410
 ; PRIOR FILING DATE: 2001-09-10
 ; PRIOR APPLICATION NUMBER: 60/261,013
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,026
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/261,029
 ; PRIOR FILING DATE: 2001-01-11
 ; PRIOR APPLICATION NUMBER: 60/313,170
 ; PRIOR FILING DATE: 2001-08-17
 ; NUMBER OF SEQ ID NOS: 306
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 110

100



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 65.3967 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-14

Perfect score: 204

Sequence: 1 MLSHLYVSKKENISKEND.....VLDEKSEAEETEEBLEEK 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Genesep_29Jan04.*
1: Genesep1980s.*
2: Genesep1990s.*
3: Genesep2000s.*
4: Genesep2001s.*
5: Genesep2002s.*
6: Genesep2003as.*
7: Genesep2003bs.*
8: Genesep2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	47.5	28	2	AAR54687
2	97	47.5	64	2	AAR54684
3	97	47.5	64	6	ABR41813
4	87	42.6	1077	4	ABR58720
5	83	40.7	31	6	ABR41855
6	83	40.7	89	4	ABG04390
7	83	40.7	106	4	ABG06955
8	83	40.7	197	6	ABP56385
9	83	40.7	379	4	ABG26552
10	83	40.7	788	4	ABG29229
11	82	40.2	31	6	ABR41854
12	82	40.2	57	4	ABG11240
13	82	40.2	62	4	ABG03970
14	82	40.2	295	4	ABG11244
15	82	40.2	412	3	AAG19532
16	82	40.2	414	3	AAG19531
17	82	40.2	456	3	AAG19530
18	82	40.2	3135	2	AAR57474
19	82	40.2	3135	3	ABR18223
20	81	39.7	81	3	AAG37760
21	81	39.7	81	3	AAG15127
22	81	39.7	81	3	AAG37763
23	81	39.7	105	3	AAG36862
24	81	39.7	105	3	AAG37759
25	81	39.7	105	3	AAG37762

26	81	39.7	105	3	AAG15126	Aag15126 Arabidops
27	81	39.7	120	3	AAG36861	Aag36861 Arabidops
28	81	39.7	120	3	AAG15125	Aag15125 Arabidops
29	81	39.7	120	3	AAG37761	Aag37761 Arabidops
30	81	39.7	120	3	AAG37758	Aag37758 Arabidops
31	81	39.7	382	4	ABG04351	Abg04351 Novel hum
32	80	39.2	31	6	ABR41853	AbR41853 Peptide u
33	80	39.2	65	4	ABG06896	Abg06896 Novel hum
34	80	39.2	85	4	ABG05351	Abg05351 Novel hum
35	80	39.2	93	4	ABG02968	Abg02968 Novel hum
36	80	39.2	93	4	ABG04398	Abg04398 Novel hum
37	80	39.2	93	4	ABG26512	Abg26512 Novel hum
38	80	39.2	98	4	ABG11735	Abg11735 Novel hum
39	80	39.2	109	4	ABG06515	Abg06515 Novel hum
40	80	39.2	109	4	ABG06958	Abg06958 Novel hum
41	80	39.2	116	4	ABG11262	Abg11262 Novel hum
42	80	39.2	226	4	ABG15564	Abg15564 Novel hum
43	80	39.2	234	4	ABG26550	Abg26550 Novel hum
44	80	39.2	1229	4	ABG24366	Abg24366 Novel hum
45	79	38.7	53	4	ABG06901	Abg06901 Novel hum

ALIGNMENTS

RESULT 1
AAR54687
ID AAR54687 standard; protein; 28 AA.
XX AAR54687;
XX
DT 25-MAR-2003 (revised)
DT 15-DEC-1994 (first entry)
XX Plasmodium falciparum merozoite 48KD surface protein fragment IV.
XX Merozoite 48KD surface antigen; immunogen; vaccine; malaria;
XX cytophilic antibody.
XX Plasmodium falciparum.
XX Plasmodium falciparum.
XX WO9409140-A1.
XX
XX 28-APR-1994.
XX
XX 18-OCT-1993; 93WO-FR001024.
XX
XX 19-OCT-1992; 92FR-00012488.
XX (INSP) INST PASTEUR.
XX Drulhe P, Bouharoun-Tayoun H, Oeuvery C;
XX WPI; 1994-151325/18.
XX N-PSDB; AAO64680.
XX New Plasmodium antigen generating cytophilic - mimicking protection
XX induced by long term exposure to the parasite, useful in vaccination
XX against and diagnosis of malaria, and prodn of antibodies for passive
XX immunisation.
XX Claim 4; Page 37; 52pp; French.
XX This sequence is a preferred fragment of the 48KD merozoite surface
XX antigen of P.falciparum which is useful as an immunogen to generate
XX antibodies (Abs). The Abs can induce a cytotoxic reaction against the
XX erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX Sequence 28 AA;
SQ
Query Match 47.5%; Score 97; DB 2; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.00065;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
 DB 9 MLSHLYVSSKDKENISKENE 28

RESULT 2
 AAR54684
 ID AAR54684 standard; protein; 64 AA.
 AC AAR54684;
 XX
 XX 25-MAR-2003 (revised)
 DT 15-DEC-1994 (first entry)
 XX
 XX Plasmodium falciparum merozoite 48KD surface protein fragment I.
 DE
 XX Merozoite 48kd surface antigen; immunogen; vaccine; malaria;
 XX cytophilic antibody.
 XX
 OS Plasmodium falciparum.
 XX
 XX WO9409140-A1.
 XX
 PD 28-APR-1994.
 XX
 PF 18-OCT-1993; 93WO-FR001024.
 XX
 PR 19-OCT-1992; 92FR-00012488.
 XX
 XX (INSP) INST PASTEUR.
 PA
 XX Drulhe P, Bouharoun-Tayoun H, Oeuvray C;
 XX WPI; 1994-151325/18.
 DR N-PSDB; AAQ64677.
 XX
 XX New Plasmodium antigen generating cytophilic - mimicking protection
 PT induced by long term exposure to the parasite, useful in vaccination
 PT against and diagnosis of malaria, and prodn of antibodies for passive
 PT immunisation.
 XX
 PS Claim 3; Page 36; 52pp; French.
 XX
 XX This sequence is a preferred fragment of the 48KD merozoite surface
 CC antigen of P.falciparum which is useful as an immunogen to generate
 CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the
 CC erythrocyte stage of the parasite. Specific subfragments of AAR54684 are
 CC separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to
 CC correct FN field.)
 XX
 SQ Sequence 64 AA;

Query Match 47.5%; Score 97; DB 2; Length 64;
 Best Local Similarity 95.0%; Pred.No. 0.0017;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
 DB 45 MLSHLYVSSKDKENISKENE 64

RESULT 3
 ABR41813
 ID ABR41813 standard; protein; 64 AA.
 AC ABR41813;
 XX
 XX 11-AUG-2003 (first entry)
 DT
 XX Merozoite surface antigen MSP-3 peptide (aa194-257).
 DE
 XX

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine.
 XX Plasmodium falciparum.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 27..37
 FT /note= "epitope specifically referred to in Claim 3"
 XX
 PN GB2378949-A.
 XX
 XX 26-FEB-2003.
 PD
 XX 16-AUG-2001; 2001GB-00020057.
 PF
 XX 16-AUG-2001; 2001GB-00020057.
 PR
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Dziegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX
 PS Disclosure; Fig 3; 75pp; English.
 XX
 CC The present sequence is that of amino acids 194-257 of the merozoite
 CC surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
 CC invention relates to the construction of recombinant human antibodies
 CC against MSP-3, in particular MSP-3(194-257) antigen. The antibodies (see
 CC also ABR41816-21) are able to passively induce naturally occurring non-
 CC sterile malaria immunity (termed premunition) and exert their effect via
 CC the antibody-dependent cellular inhibition of growth (ADCI) mechanism.
 CC They can be used in the diagnosis and treatment of malaria. The invention
 CC also provides a method of screening a phage library with target
 CC molecules, such as MSP-3 antigen, bound to beads, preferably magnetic
 CC microspheres
 XX
 SQ Sequence 64 AA;

Query Match 47.5%; Score 97; DB 6; Length 64;
 Best Local Similarity 95.0%; Pred.No. 0.0017;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
 DB 45 MLSHLYVSSKDKENISKENE 64

RESULT 4
 ABB58720
 ID ABB58720 standard; protein; 1077 AA.
 XX
 AC ABB58720;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 2952.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW Drosophila melanogaster.
 XX
 OS WO200171042-A2.
 XX
 PN 27-SEP-2001.
 XX
 PD
 XX

Claim 20; SEQ ID NO 34749; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 89 AA;

Query Match 40.7%; Score 83; DB 4; Length 89;
 Best Local Similarity 51.4%; Pred. No. 0.086;
 Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 7 VSKKKNISKENDVDLDEKEEAETEELEEK 41
 ||||| : : : : : ||||| : : : : :
 Db 41 VSKKKKKKKKEEEEEEEEEEEEEEEEEEE 75

RESULT 7
 ABG06955
 ID ABG06955 standard; protein; 106 AA.

XX AC ABG06955;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #6946.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX QY Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS71142.

XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX PS Claim 20; SEQ ID NO 37314; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain

CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX
 SQ Sequence 106 AA;

Query Match 40.7%; Score 83; DB 4; Length 106;
 Best Local Similarity 48.6%; Pred. No. 0.11;
 Matches 17; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 7 VSKKKNISKENDVDLDEKEEAETEELEEK 41
 ||||| : : : : : ||||| : : : : :
 Db 41 VSEKEEEDDEEEEEEEEEEEEEEEEEEE 75

RESULT 8

ABP56385

ID ABP56385 standard; peptide; 197 AA.

XX AC ABP56385;

XX DT 11-MAR-2003 (first entry)

XX DE Polyanionic fusion protein product #5.

XX KW Polyanionic polymer; bioactivity; water solubility.

XX OS Synthetic.

XX PN WO200277036-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US008614.

XX PR 21-MAR-2001; 2001US-0277705P.

XX PA (LEUN/) LEUNG D W.

XX PI Leung DW, Bergman PA, Lofquist A, Pietz GE, Tompkins CK;

XX PI Waggoner DW;

XX DR WPI; 2003-058367/05.

XX PT Producing monodispersed preparation of polyanionic polymer for therapy,
 PT by expressing vector comprising ligation product of oligonucleotides
 PT encoding glutamate/aspartate residues in host cell and isolating the
 PT product.

XX PS Example 7; Page 46; 74pp; English.

XX CC The present invention describes a method (M) for producing a
 CC monodispersed preparation of a polyanionic polymer (PP) larger than 10
 CC kD. (M) involves inserting into an expression vector (EV) a ligation
 CC product formed by ligating together oligonucleotides that encode
 CC glutamate/aspartate residues, expressing EV in a host cell, and isolating
 CC the protein product (P) of EV, where (P) is PP and at least 80% of PP is

CC approximately of the same molecular weight. Also described: (1) a
CC recombinant fusion protein (I) comprising a polyanionic polypeptide and
CC another polypeptide at either one end or at both ends of it; (2) a
CC polyanionic polymer (II) conjugate comprising a polyanionic polymer and
CC leukine, where the polyanionic polymer is polyglutamic acid or
CC polyaspartic acid; (3) a vector (III) comprising a cassette which
CC comprises a nucleotide sequence encoding a polyanionic polymer and at
CC least one other nucleotide sequence, where the polyanionic polymer is
CC polyglutamic acid or polyaspartic acid; (4) production of (I); (5) a cell
CC (IV) comprising (III) or a vector that comprises a nucleotide sequence
CC that encodes a polyanionic polymer that is larger than 10 kDa; and (6) a
CC recombinantly-produced polyanionic polymer (V) that is of any molecular
CC weight or is larger than 10 kDa, and is conjugated to another protein. (I)
CC is useful for treating a disease or ailment in an individual by
CC administering (I) to the individual. (I) is also useful for delivering an
CC effective amount of a pharmaceutically active agent, a therapeutic
CC protein or a drug to a patient in need of it, or for diagnostic and
CC testing or research purposes. AB222045 to AB222111 and ABP6374 to
CC ABP56400 represent sequences used in the exemplification of the present
CC invention

XX Sequence 197 AA;

Query Match 40.7%; Score 83; DB 6; Length 197;
Best Local Similarity 41.5%; Pred. No. 0.22;
Matches 17; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
DB 10 MPENLYFGHEEEEEEEEEEEEEEEEEEEEEEE 50

RESULT 9

ABG26552
ID ABG26552 standard; protein; 379 AA.

XX AC ABG26552;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #26543.

XX XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS90739.

XX XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 56911; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC in diagnostics as expressed sequence tags for identifying expressed

CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX CC
XX SQ Sequence 379 AA;

Query Match 40.7%; Score 83; DB 4; Length 379;
Best Local Similarity 51.4%; Pred. No. 0.46;
Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

QY 7 VSSKDKENISKENDVLDKEKEAEETEEELK 41
DB 13 VSKKXKXKKEEEEEEEEEEEEEEEEEEE 47

RESULT 10

ABG29229
ID ABG29229 standard; protein; 788 AA.

XX AC ABG29229;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #29220.

XX XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS93416.

XX XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

XX PS Claim 20; SEQ ID NO 59588; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 40.7%; Score 83; DB 4; Length 788;
Best Local Similarity 51.4%; Pred. No. 1.1;
Matches 18; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Sequence 788 AA;
1 VSKKKENISKENDVDLDEKEEAETEELEEK 41
427 VSKKKKKKKKKKKKKKKKKKKKKKKKKKK 461

RESULT 11
ABR41854
ID ABR41854 standard; peptide; 31 AA.
XX ABR41854;
XX
DT 11-AUG-2003 (first entry)
XX
DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
KW immunotherapy; vaccine; epitope.
XX
OS Plasmodium falciparum.
XX
FH Key Location/Qualifiers
FT Misc-difference 15..16 /note="linker residues"
XX
H GB2378949-A.
XX
PD 26-FEB-2003.
XX
PF 16-AUG-2001; 2001GB-00020057.
XX
PR 16-AUG-2001; 2001GB-00020057.
XX
PA (DZIE/) DZIEGIEL M S H.
PA (LUND/) LUNDQUIST R.
PA (NIEL/) NIELSEN L K.
XX
PI Dzigiel MSH, Lundquist R, Nielsen LK;
XX
DR WPI; 2003-303108/30.
XX
PT Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues

15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;
Query Match 40.2%; Score 82; DB 6; Length 31;
Best Local Similarity 89.5%; Pred. No. 0.033;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MLSHLYVSSKKENISKEN 19
13 MLXXLYVSSKKENISKEN 31

RESULT 12
ABG11240
ID ABG11240 standard; protein; 57 AA.
XX
AC ABG11240;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #11231.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS75427.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 41599; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (II) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The

PS
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (I) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and

Search completed: September 22, 2004, 18:06:24
Job time : 68.3967 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 19.314 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1 MLSHLYVSKDKENISKEND.....VLDEKSEAEETEEELK 41

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	47.5	28	3	US-08-416-711-4
2	97	47.5	28	4	US-09-356-497-4
3	97	47.5	64	3	US-08-416-711-1
4	97	47.5	64	4	US-09-356-497-1
5	82	40.2	3135	1	US-08-323-1708-2
6	82	40.2	3135	4	US-08-954-441-2
7	79	38.7	714	2	US-08-990-114-3
8	79	38.7	714	4	US-09-241-333-3
9	77	37.7	740	1	US-08-257-073-5
10	76	37.3	905	2	US-08-574-959A-9
11	76	37.3	905	3	US-09-357-014-9
12	76	37.3	1135	2	US-08-574-959A-7
13	76	37.3	1135	3	US-09-357-014-7
14	75	36.8	739	3	US-09-022-983-2
15	73	35.8	111	4	US-09-886-319A-9
16	73	35.8	111	4	US-09-886-319A-10
17	73	35.8	1525	4	US-09-418-710-69
18	73	35.8	1912	4	US-09-495-714C-2
19	73	35.8	1977	4	US-09-495-714C-4
20	71.5	35.0	1312	4	US-09-345-882-29
21	70	34.3	66	4	US-09-621-976-5084
22	70	34.3	77	4	US-09-621-976-7521
23	70	34.3	1162	2	US-08-728-323A-2
24	70	34.3	1162	4	US-09-298-568-2
25	70	34.3	1162	4	US-09-410-399-2
26	69.5	34.1	357	1	US-08-552-142A-2
27	69.5	34.1	357	1	US-08-910-973-2

Sequence 2, Appli
Sequence 2, Appli
Sequence 27, Appli
Sequence 29, Appli
Sequence 24, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 5087, Ap
Sequence 5083, Ap
Sequence 29, Appli
Sequence 4, Appli
Sequence 7, Appli
Sequence 214, App
Sequence 8, Appli
Sequence 14, Appli
Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-416-711-4
; Sequence 4, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; APPLICANT: OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-4

Query Match 47.5%; Score 97; DB 3; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.00016;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28

RESULT 2
US-09-356-497-4
; Sequence 4, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-356-497-4

Query Match 47.5%; Score 97; DB 4; Length 28;
Best Local Similarity 95.0%; Pred. No. 0.00016;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
Db 9 MLSHLYVSSKDKENISKENE 28

RESULT 3
US-08-416-711-1
; Sequence 1, Application US/08416711

Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-1

Query Match 47.5%; Score 97; DB 3; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.00039;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKEND 20
Db 45 MLSHLYVSSKDKENISKENE 64

RESULT 4
US-09-356-497-1
; Sequence 1, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

;; COUNTRY: USA
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/356.497
;; FILING DATE: 19-Jul-1999
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/416,711
;; FILING DATE: 08-AUG-1995
;; APPLICATION NUMBER: PCT/FR93/01024
;; FILING DATE: 18-OCT-1993
;; APPLICATION NUMBER: FR 92/12488
;; FILING DATE: 19-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 64 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 47.5%; Score 97; DB 4; Length 64;
Best Local Similarity 95.0%; Pred. No. 0.00039;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSLHYVSSKDKENISKEND 20
Db 45 MSLHYVSSKDKENISKENE 64

RESULT 5
US-08-323-170B-2
; Sequence 2, Application US/08323170B
; Patent No. 5733772
; GENERAL INFORMATION:
; APPLICANT: Williamson, Kim C.
; APPLICANT: Kaelow, David C.
; TITLE OF INVENTION: Cloning and Expression of Plasmodium
; TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pfs230
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,170B
; FILING DATE: 13-OCT-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,409
; FILING DATE: 29-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 015280-113110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-441-2

Query Match 40.2%; Score 82; DB 4; Length 3135;
Best Local Similarity 51.6%; Pred. No. 1.2;

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Quine, Jonathan A.
;; REGISTRATION NUMBER: P-41,261
;; REFERENCE/DOCKET NUMBER: 015280-113100US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3135 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-323-170B-2

Query Match 40.2%; Score 82; DB 1; Length 3135;
Best Local Similarity 51.6%; Pred. No. 1.2;
Matches 16; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 11 DKENISKENDVLDKEKEEAETEELEEK 41
Db 265 DEEDMSPRDNFVIDDEEEEEEEEEEE 295

RESULT 6
US-08-954-441-2
; Sequence 2, Application US/08954441
; Patent No. 6316000
; GENERAL INFORMATION:
; APPLICANT: Williamson, Kim C.
; APPLICANT: Kaelow, David C.
; TITLE OF INVENTION: Cloning and Expression of Plasmodium
; TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pfs230
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,441
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/323,170
; FILING DATE: 13-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/010,409
; FILING DATE: 29-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 015280-113110US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0300
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-441-2

; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
; FILING DATE: 11-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/852,305
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,183
; FILING DATE: 20-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTWS
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 740 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-257-073-5

Query Match 37.7%; Score 77; DB 1; Length 740;
Best Local Similarity 50.0%; Pred. No. 0.86;
Matches 16; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 10 KDKENISKENDVLDKEKEAEETEELEEK 41
Db 695 KEKEKEKEKEKEKEKEKEKEKEKEE 726

RESULT 10
US-08-574-959A-9
; Sequence 9, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

Query Match 37.3%; Score 76; DB 3; Length 905;
Best Local Similarity 40.0%; Pred. No. 1.4;
Matches 16; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 2 LSHLYVSSKDKENISKENDVLDKEKEAEETEELEEK 41
Db 644 LTVININSDEEEEEEEEEEEEEEEEEEE 683

RESULT 12
US-08-574-959A-7
; Sequence 7, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:

; LENGTH: 905 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-574-959A-9

Query Match 37.3%; Score 76; DB 2; Length 905;
Best Local Similarity 40.0%; Pred. No. 1.4;
Matches 16; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 2 LSHLYVSSKDKENISKENDVLDKEKEAEETEELEEK 41
Db 644 LTVININSDEEEEEEEEEEEEEEEEEEE 683

RESULT 11
US-09-357-014-9
; Sequence 9, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/357,014
; FILING DATE: 19-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/574,959
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; MOLECULE TYPE: protein
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-357-014-9

Query Match 37.3%; Score 76; DB 3; Length 905;
Best Local Similarity 40.0%; Pred. No. 1.4;
Matches 16; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

QY 2 LSHLYVSSKDKENISKENDVLDKEKEAEETEELEEK 41
Db 644 LTVININSDEEEEEEEEEEEEEEEEEEE 683

RESULT 12
US-08-574-959A-7
; Sequence 7, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:

APPLICANT: Jaekyoon Shin, Insil Joong, Ratna K. Vadlamudi
 APPLICANT: and Jack L. Strominger
 TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
 TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
 TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, Suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/574,959A
 FILING DATE: 19-DEC-95
 ATTORNEY/AGENT INFORMATION:
 NAME: Mandragouras, Amy E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: DFN-008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1135 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-574-959A-7

Query Match 37.3%; Score 76; DB 2; Length 1135;
 Best Local Similarity 40.0%; Pred. No. 1.8;
 Matches 16; Conservative 11; Mismatches 13; Indels

QY 2 LSHLYVSSKDKENISKENDDDVLDEKEEAAETEEEELEEK 41
| : : | | : | : : : | | | | : :
Db 874 LTVININSSDEEEEEEGEEEEEEEEEEEEEEEEEEEE 913

RESULT 13
US-09-357-014-7
Sequence 7, Application US/09357014
Patent No. 6291645
GENERAL INFORMATION:
APPLICANT: Jaekyoon Shin, Inail Joung, Ratna K. Vadlamudi
and Jack L. Strominger
TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
AND USES THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/357,014
FILING DATE: 19-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/574,959
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:

```

; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
; US-09-357-014-7

Query Match 37.3%; Score 76; DB 3; Length 1135;
Best Local Similarity 40.0%; Pred No. 1.8;
Matches 16; Conservative 11; Mismatches 13; Indels 0

QY 2 LSHLYVSKDKENISKENDVLDEKEEAEETEEELK 41
   | : : | | : : : : : : | | | | | :
DB 874 LTVINSSDEEEEGEEEEEEEEEEEEEE 913

RESULT 14
US-09-022-983-2
; Sequence 2, Application US/09022983
; Patent No. 6159731
; GENERAL INFORMATION:
; APPLICANT: Yang, Xiaolu
; APPLICANT: Khosravi-Far, Roya
; APPLICANT: Chang, Howard Y.
; APPLICANT: Baltimore, David
; TITLE OF INVENTION: DAXX, A NOVEL FAS-BINDING
; TITLE OF INVENTION: PROTEIN THAT ACTIVATES JNK AND APOPTOSIS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/022,983
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,919
; FILING DATE: 12-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/051,753
; FILING DATE: 26-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Van Amsterdam, John R.
; REGISTRATION NUMBER: 40,212
; REFERENCE/DOCKET NUMBER: M0656/7036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

```

US-09-022-983-2

Query Match 36.8%; Score 75; DB 3; Length 739;
 Best Local Similarity 50.0%; Pred. No. 1.4;
 Matches 17; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
 Qy 8 SSKDKENISKENDVLDKEKEAEETEEELK 41
 Db 436 TSKAETDDDDDDDDNESEEEEEEEK 469

RESULT 15

US-09-886-319A-9
 ; Sequence 9, Application US/09886319A
 ; Patent No. 6586185
 ; GENERAL INFORMATION:
 ; APPLICANT: Wolf, Eckard
 ; APPLICANT: Werner, Sabine
 ; APPLICANT: Halle, Jörn-Peter
 ; APPLICANT: Regenbogen, Johannes
 ; APPLICANT: Goppelt, Andreas
 ; TITLE OF INVENTION: Use of Polypeptides or Nucleic Acids for
 ; TITLE OF INVENTION: the Diagnosis or Treatment of Skin Disorders and Wound
 ; TITLE OF INVENTION: Healing and for the Identification of Pharmacologically
 ; TITLE OF INVENTION: Active Substances
 ; FILE REFERENCE: 50125/014002
 ; CURRENT APPLICATION NUMBER: US/09/886,319A
 ; CURRENT FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: US 60/222,081
 ; PRIOR FILING DATE: 2000-08-01
 ; PRIOR APPLICATION NUMBER: DE 10030149.5
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 9
 ; LENGTH: 111
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-886-319A-9

Query Match 35.8%; Score 73; DB 4; Length 111;
 Best Local Similarity 46.9%; Pred. No. 0.3;
 Matches 15; Conservative 9; Mismatches 8; Indels 0; Gaps 0;
 Qy 8 SSKDKENISKENDVLDKEKEAEETEEELK 39
 Db 38 NAQNEENGQADNEVDDEEEEGCGEEEEEE 69

Job completed: September 22, 2004, 18:11:03
 Time : 20.314 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 13.1901 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKKENMLSHLYVSSKDKENISKEND 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	97.3	28	3	US-08-416-711-4
2	143	97.3	28	4	US-09-356-497-4
3	143	97.3	64	3	US-08-416-711-1
4	143	97.3	64	4	US-09-356-497-1
5	48.5	33.0	652	1	US-08-261-663A-6
6	48.5	33.0	652	4	US-09-357-206A-5
7	48.5	33.0	652	4	US-09-813-742A-5
8	48.5	33.0	652	5	PCT-US95-07754A-6
9	48.5	33.0	746	4	US-09-134-001C-3214
10	47.5	32.3	68	4	US-09-621-976-7228
11	47.5	32.3	68	4	US-09-621-976-7229
12	47.5	32.3	68	4	US-09-621-976-7234
13	47.5	32.3	889	1	US-08-118-101A-4
14	47.5	32.3	1073	3	US-09-541-782-6
15	47.5	32.3	1073	4	US-09-723-820-6
16	47.5	32.3	1073	4	US-10-270-085-6
17	47.5	32.3	1177	4	US-09-134-001C-5106
18	47.5	32.3	10182	4	US-09-134-001C-3159
19	47	32.0	534	4	US-09-508-370A-5
20	46	31.3	27	3	US-08-416-711-3
21	46	31.3	27	4	US-09-356-497-3
22	46	31.3	181	4	US-09-134-001C-3897
23	46	31.3	516	4	US-09-496-320-13
24	45.5	31.0	239	4	US-09-513-783A-50
25	45.5	31.0	514	4	US-09-198-452A-214
26	45.5	31.0	541	4	US-09-513-783A-34
27	45.5	31.0	917	4	US-09-049-698-41

Sequence 6, Appli
Sequence 6428, Ap
Sequence 7850, Ap
Sequence 4697, Ap
Sequence 9, Appli
Sequence 4201, Ap
Sequence 5, Appli
Sequence 5, Appli
Sequence 19, Appli
Sequence 4, Appli
Sequence 4, Appli
Sequence 11, Appli
Sequence 8, Appli
Sequence 3818, Ap
Sequence 7803, Ap
Sequence 1045, Ap
Sequence 168, App

28 45.5 31.0 2184 4 US-09-417-485D-6
29 45 30.6 88 4 US-09-621-976-6428
30 45 30.6 162 4 US-09-489-039A-7850
31 45 30.6 208 4 US-09-134-001C-4697
32 45 30.6 308 1 US-07-828-700-9
33 45 30.6 384 4 US-09-134-001C-4201
34 45 30.6 487 4 US-09-351-224E-5
35 45 30.6 487 4 US-09-677-488A-5
36 45 30.6 487 4 US-09-677-488B-5
37 45 30.6 523 2 US-08-482-728A-19
38 45 30.6 523 3 US-09-028-366-4
39 45 30.6 523 4 US-09-715-285-4
40 45 30.6 642 4 US-09-371-338-11
41 45 30.6 734 4 US-09-585-858-8
42 44.5 30.3 60 4 US-09-134-001C-3818
43 44.5 30.3 279 4 US-09-543-681A-7803
44 44 29.9 52 4 US-09-205-258-1045
45 44 29.9 111 4 US-09-247-155-168

ALIGNMENTS

RESULT 1

US-08-416-711-4
; Sequence 4, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; APPLICANT: OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-4

```

Query Match          97.3%; Score 143; DB 3; Length 28;
Best Local Similarity 96.4%; Pred. No. 9.2e-15;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKGENMLSHLYVSSKDKENISKEND 28
   |||||
Db 1 PEHKGENMLSHLYVSSKDKENISKENE 28

RESULT 2
US-09-356-497-4
; Sequence 4, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
;              BOUHAROUN-TAYOUN, HASNAQ
;              OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
;              PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;              P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
;
US-09-356-497-4

Query Match          97.3%; Score 143; DB 4; Length 28;
Best Local Similarity 96.4%; Pred. No. 9.2e-15;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKGENMLSHLYVSSKDKENISKEND 28
   |||||
Db 1 PEHKGENMLSHLYVSSKDKENISKENE 28

RESULT 3
US-08-416-711-1
; Sequence 1, Application US/08416711

```

```

; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; APPLICANT: OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-416-711-1

Query Match 97.3%; Score 143; DB 3; Length 64;
Best Local Similarity 96.4%; Pred. No. 2.5e-14;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 37 PEHKKEENMLSHLYVSSKDKENISKENE 64
|||||||||||||||||||||||||
|||||||||||||||||||||||||

RESULT 4
US-09-356-497-1
; Sequence 1, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

```

RESULT 3
US-08-416-711-1
; Sequence 1, Application US/08416711

COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/356.497
FILING DATE: 19-Jul-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/416,711
FILING DATE: 08-AUG-1995
APPLICATION NUMBER: PCT/FR93/01024
FILING DATE: 18-OCT-1993
APPLICATION NUMBER: FR 92/12488
FILING DATE: 19-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-085-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 64 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 97.3%; Score 143; DB 4; Length 64;
Best Local Similarity 96.4%; Pred. No. 2.5e-14;
Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKEENMLSHLVSSKDKENISKEND 28
DB 37 PEHKEENMLSHLVSSKDKENISKENE 64

RESULT 5
US-08-261-663A-6
Sequence 6, Application US/08261663A
Patent No. 5571706
GENERAL INFORMATION:
APPLICANT: Baker, Barbara J
TITLE OF INVENTION: Plant Virus Resistance Gene and Methods
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Margaret A. Connor, USDA-ARS
STREET: 800 Buchanan Street
CITY: Albany
STATE: CA
COUNTRY: USA
ZIP: 94710
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261.663A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Connor, Margaret A
REGISTRATION NUMBER: 30043
REFERENCE/DOCKET NUMBER: 0094.94

TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 559-6067
TELEFAX: (510) 559-5777
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 652 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-261-663A-6

Query Match 33.0%; Score 48.5; DB 1; Length 652;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 8 NMLSHLVSSKDKENIS-KEND 28
DB 606 NSLRHLWTETKKNIAKEGD 627

RESULT 6
US-09-357-206A-5
Sequence 5, Application US/09357206A
Patent No. 6372962
GENERAL INFORMATION:
APPLICANT: Dinesh-Kumar, S.
TITLE OF INVENTION: Pathogen Resistance in Plants using cdNA-N/Intron Constructs
FILE REFERENCE: 042250/191805 (5830-5)
CURRENT APPLICATION NUMBER: US/09/357,206A
CURRENT FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: US 60/093,494
PRIOR FILING DATE: 1998-07-20
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 5
LENGTH: 652
TYPE: PRT
ORGANISM: Nicotiana glutinosa
US-09-357-206A-5

Query Match 33.0%; Score 48.5; DB 4; Length 652;
Best Local Similarity 50.0%; Pred. No. 48;
Matches 11; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 8 NMLSHLVSSKDKENIS-KEND 28
DB 606 NSLRHLWTETKKNIAKEGD 627

RESULT 7
US-09-813-742A-5
Sequence 5, Application US/09813742A
Patent No. 6630618
GENERAL INFORMATION:
APPLICANT: Baker, Barbara J
TITLE OF INVENTION: NON-PATHOGEN INDUCED SYSTEMIC ACQUIRED RESISTANCE (SAR) IN PLANTS
FILE REFERENCE: 42250/209601 (5830-12)
CURRENT APPLICATION NUMBER: US/09/813,742A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,027
PRIOR FILING DATE: 2000-03-21
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patentin version 3.1
SEQ ID NO 5
LENGTH: 652
TYPE: PRT
ORGANISM: Nicotiana glutinosa
US-09-813-742A-5

Query Match 33.0%; Score 48.5; DB 4; Length 652;
Best Local Similarity 50.0%; Pred. No. 48;

```

RESULT 9
US-09-134-001C-3214
; Sequence 3214, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3214
; LENGTH: 746

```

```

Qy 6 ENMLSHLYVSSKDKENISKEND 28
   ||| ||| : ||| : |||
Db 685 EQOSLG-LYDSGQDENITHEDD 706

RESULT 14
US-09-541-782-6
; Sequence 6, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541,782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-541-782-6

Query Match 32.3%; Score 47.5; DB 3; Length 1073;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

Qy 2 EHKKE-ENMLSHLYVSSKDKENISKEND 28
   ||||| : ||| : ||| : |||
Db 499 EHKKEVALQLVNSSTELSVKSENE 526

RESULT 15
US-09-723-820-6
; Sequence 6, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723,820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1073
; TYPE: PRT
; ORGANISM: S.pombe
US-09-723-820-6

Query Match 32.3%; Score 47.5; DB 4; Length 1073;
Best Local Similarity 42.9%; Pred. No. 1.2e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 1; Gaps 1;

Qy 2 EHKKE-ENMLSHLYVSSKDKENISKEND 28
   ||||| : ||| : ||| : |||
Db 499 EHKKEVALQLVNSSTELSVKSENE 526

Search completed: September 22, 2004, 18:11:02
Job time : 14.1901 secs

```



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:09:13 ; Search time 60.3967 Seconds
(without alignments)
148.878 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKKEENMLSHLYVSSKOKENISKEND 28

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	28	14	US-10-294-770-13
2	147	100.0	28	16	US-10-774-602-13
3	143	97.3	28	14	US-10-294-770-4
4	143	97.3	28	16	US-10-238-741-4
5	143	97.3	28	16	US-10-774-602-4
6	143	97.3	64	14	US-10-294-770-1
7	143	97.3	64	16	US-10-238-741-1
8	143	97.3	64	16	US-10-774-602-1
9	101	68.7	41	14	US-10-294-770-14
10	101	68.7	41	16	US-10-774-602-14
11	54	36.1	199	16	US-10-437-963-161536
12	53	36.1	92	16	US-10-437-963-109657
13	53	36.1	426	9	US-09-731-872-310
14	53	36.1	426	9	US-09-731-872-317
15	53	36.1	426	10	US-09-876-997-310

16	53	36.1	426	10	US-09-876-997-317	Sequence 317, App
17	53	36.1	426	16	US-10-655-601-5	Sequence 5, Appli
18	53	36.1	481	9	US-09-731-872-415	Sequence 415, App
19	53	36.1	481	10	US-09-876-997-415	Sequence 415, App
20	52.5	35.7	307	16	US-10-437-963-140346	Sequence 140346, A
21	52.5	35.7	654	12	US-10-425-114-65105	Sequence 65105, A
22	52	35.4	81	12	US-10-424-599-150366	Sequence 150366, A
23	52	35.4	443	12	US-10-282-122A-48757	Sequence 48757, A
24	52	35.4	749	15	US-10-369-493-21933	Sequence 21933, A
25	51.5	35.0	1182	12	US-10-282-122A-53445	Sequence 53445, A
26	51	34.7	98	16	US-10-437-963-159924	Sequence 159924, A
27	51	34.7	174	14	US-10-238-075-558	Sequence 558, App
28	51	34.7	204	16	US-10-767-701-31636	Sequence 31636, A
29	51	34.7	415	12	US-10-282-122A-70442	Sequence 70442, A
30	51	34.7	572	12	US-10-267-503-235	Sequence 235, App
31	51	34.7	572	16	US-10-287-226-661	Sequence 661, App
32	51	34.7	1441	16	US-10-437-963-159919	Sequence 159919, A
33	50.5	34.4	258	12	US-10-424-599-246340	Sequence 246340, A
34	50	34.0	515	12	US-10-282-122A-53329	Sequence 53329, A
35	49	33.3	209	16	US-10-437-963-117713	Sequence 117713, A
36	49	33.3	231	12	US-10-425-114-68769	Sequence 68769, A
37	49	33.3	344	12	US-10-425-114-45399	Sequence 45399, A
38	49	33.3	405	12	US-10-425-114-68612	Sequence 68612, A
39	49	33.3	473	15	US-10-094-749-1936	Sequence 1936, App
40	49	33.3	473	16	US-10-408-765A-2491	Sequence 2491, App
41	49	33.3	680	16	US-10-437-963-123028	Sequence 123028, A
42	49	33.3	741	16	US-10-437-963-117706	Sequence 117706, A
43	48.5	33.0	652	9	US-09-813-742-5	Sequence 5, Appli
44	48.5	33.0	743	12	US-10-282-122A-70586	Sequence 70586, A
45	48.5	33.0	1700	16	US-10-437-963-162893	Sequence 162893, A

ALIGNMENTS

RESULT 1
US-10-294-770-13
; Sequence 13, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUIHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-294-770-13

Query Match 100.0%; Score 147; DB 14; Length 28;

Best Local Similarity 100.0%; Pred. No. 1,1e-12;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKOKENISKEND 28

Db 1 PEHKKEENMLSHLYVSSKOKENISKEND 28

RESULT 2

US-10-774-602-13

; Sequence 13, Application US/10774602

```

; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBOLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238.741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356.497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416.711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBOLON, NORMAN F.
; REGISTRATION NUMBER: 24.618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-238-741-4
Query Match 97.3%; Score 143; DB 16; Length 28;
Best Local Similarity 96.4%; Pred.No.3.9e-12;
Matches 27; Conservative 1; Mismatches 0; Indels 0;
QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
DB 1 PEHKKEENMLSHLYVSSKDKENISKENE 28
RESULT 5
US-10-774-602-4
; Sequence 4, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTEC
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774.602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356.947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238.741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416.711
; PRIOR FILING DATE: 1995-08-08

```

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28

Search completed: September 22, 2004, 18:34:05
Job time : 61.3967 secs

RESULT 14
US-09-731-872-317
; Sequence 317, Application US/09731872
; Patent No. US20020103604A1

```

RESULT 14
US-09-731-872-317
; Sequence 317, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN cDNAs ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78.US3.REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; PRIORITY FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIORITY FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIORITY FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 317
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -28..-1
US-09-731-872-317

Query Match 36.1%; Score 53; DB 9; Length 426;
Best Local Similarity 33.3%; Pred. NO. 86;
Matches 13; Conservative 5; Mismatches 7; Indels 14; Gaps 1;

QY 2 EHKKEENMLSH-----LYVSSKKKENISKE 26
DB 134 QHRKEEMKMKLEEFWLNKSEVLYYTVKKGNISSQ 172

```

Miss Page Blank (1891)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:02:42 ; Search time 9.50413 Seconds
(without alignments)
253.026 Million cell updates/sec

Title: US-10-774-602-11

Perfect score: 124

Sequence: 1 YEKAKNAYQKQANQAVLKAKEASYD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	47.6	890	2	S22452
2	55.5	44.8	820	2	T51510
3	55	44.4	752	2	H6770
4	55	44.4	831	2	G41662
5	54.5	44.0	1162	2	T51040
6	53	42.7	463	2	T13425
7	52	41.9	213	2	T22984
8	52	41.9	420	2	T48234
9	50	40.3	106	2	B82363
10	49.5	39.9	347	2	B85432
11	49	39.5	180	2	F97149
12	49	39.5	651	2	T15624
13	48	38.7	988	2	A24341
14	47.5	38.3	100	1	R5EC23
15	47.5	38.3	100	2	G91151
16	47.5	38.3	100	2	C85997
17	47.5	38.3	174	2	G70320
18	47.5	38.3	1131	2	A49393
19	47	37.9	51	2	F97708
20	47	37.9	67	2	T14410
21	47	37.9	119	2	B81939
22	47	37.9	119	2	F81165
23	47	37.9	179	2	T42293
24	47	37.9	441	2	B96953
25	47	37.9	1240	2	T04833
26	47	37.9	1386	2	AC1533
27	46.5	37.5	100	2	AC1006
28	46.5	37.5	482	2	AF1120
29	46	37.1	338	2	E69964

leucyl aminopeptid
uncharacterized ph
probable late embr
conserved hypothe
conserved hypothe
pollen coat protei
hypotheical prote
late embryogenesis
hypotheical prote
probable electron
ORF MSV194 ALI mot
mitochondrial impo
hypotheical prote
fibronectin bindin
hypotheical prote
hypotheical prote

ALIGNMENTS

RESULT 1

S22452

surface exclusion protein seal precursor - Enterococcus faecalis plasmid pAD1
C;Species: Enterococcus faecalis

C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 15-Oct-1999

C;Accession: S22452; S22069

R;Weidlich, G.; Wirth, R.; Galli, D.

Mol. Gen. Genet. 233, 161-168, 1992

A;Title: Sex pheromone plasmid pAD1-encoded surface exclusion protein of Enterococcus faecalis

A;Reference number: S22451; MUID:922931110; PMID:1603060

A;Accession: S22452

A;Molecule type: DNA

A;Residues: 1-890 <WEI>

A;Cross-references: EMBL:X62658; NID:G49019; PIDN:CAA44526.1; PID:949021

C;Genetics:

A;Gene: seal

A;Genome: plasmid

Query Match 47.6%; Score 59; DB 2; Length 890;

Best Local Similarity 61.9%; Pred. No. 3;

Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKQANQAVLKAKEA 21

||||| ||| ||| ||| |||

Db 709 YEKAVNELNKAERAVVQAKEA 729

RESULT 2

T51510

hypotheical protein F5E19 120 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000

R;Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew

submitted to the Protein Sequence Database, August 2000

C;Accession: T51510

A;Reference number: Z25394

A;Accession: T51510

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-820 <SAT>

A;Cross-references: EMBL:AL391147

A;Experimental source: cultivar Columbia; BAC clone F5E19

C;Genetics:

A;Map position: 5

A;Introns: 184/2; 230/3; 245/1; 283/1; 315/2; 337/3; 357/3; 561/1; 693/3; 767/2; 795/3

A;Note: F5E19_120

Query Match 44.8%; Score 55.5; DB 2; Length 820;

Best Local Similarity 59.1%; Pred. No. 8.4;

Matches 13; Conservative 2; Mismatches 4; Indels 3; Gaps 1;

QY 1 YEKAKNAYQKQANQAVLKAKEAS 22

```
Db          446 YEKSNAYQ---EAIKAEAS 464
||||| ||||| :|: |||||
Query Match      44.4%; Score 55; DB 2; Length 752;
Best Local Similarity 52.6%; Pred. No. 9.1;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

RESULT 3
H86770
hypothetical protein yliB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: H86770
R:Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: H86770
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-752 <STO>
A:Cross-references: GB:AE005176; PID:gl2724133; PIDN:AAK05266.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: YliB

Query Match      44.4%; Score 55; DB 2; Length 752;
Best Local Similarity 52.6%; Pred. No. 9.1;
Matches 10; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY          2 EKAKNAYQKANOAVLKAKE 20
||||| ||||| :|: |||||
Db          450 EEAKNTYHFANEMILSKK 468
||||| ||||| :|: |||||

RESULT 4
130K surface exclusion protein Sec10 precursor - Enterococcus faecalis plasmid pCF10
C:Species: Enterococcus faecalis
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 15-Oct-1999
C:Accession: G41662
R:Xao, S.M.; Olmsted, S.B.; Viksins, A.S.; Gallo, J.C.; Dunny, G.M.
J. Bacteriol. 173, 7650-7664, 1991
A:Title: Molecular and genetic analysis of a region of plasmid pCF10 containing positive
terococcus faecalis.
A:Reference number: A41662; MUID:92041679; PMID:1938961
A:Accession: G41662
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-891 <KAO>
A:Cross-references: GB:M64978; NID:gi50552; PIDN:AAA65847.1; PID:gi50554
C:Genetics:
A:Genome: plasmid

Query Match      44.4%; Score 55; DB 2; Length 891;
Best Local Similarity 57.1%; Pred. No. 11;
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY          1 YEKKNAYQKANOAVLKAKEA 21
||||| ||||| :|: |||||
Db          709 YEKALNELNKAAEAAVVOAQA 729
||||| ||||| :|: |||||

RESULT 5
T51040
hypothetical protein B15120.100 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51040
R:Schulte, U.; Algn, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51040
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1162 <SCH>
```

```
A:Cross-references: EMBL:AL389900; GSPDB:GN00116; NCSP:B15120.100
A:Experimental source: BAC clone B15120; strain OR74A
C:Genetics:
A:Gene: NCSP:B15120.100
A:Map position: 6
A:Introns: 59/2; 135/3; 171/3; 309/1; 430/2

Query Match      44.0%; Score 54.5; DB 2; Length 1162;
Best Local Similarity 59.3%; Pred. No. 16;
Matches 16; Conservative 3; Mismatches 3; Indels 5; Gaps 2;

QY          2 EKA-KNAYQ----KANQAVLKAKEASS 23
||||| ||||| :|: |||||
Db          632 EKALNKAQEMARIKAEQAALKAKEANA 658
||||| ||||| :|: |||||

RESULT 6
T13425
regulatory protein K10, oocyte-specific - fruit fly (Drosophila melanogaster)
N:Alternate names: protein EG.30B8.5
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 15-Oct-1999
C:Accession: T13425; A28826
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17668
A:Accession: T13425
A>Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-463 <MUR>
A:Cross-references: EMBL:AL009195; NID:el355203; PID:el202207; PIDN:CAA15702.1
Genes Dev. 2, 891-900, 1988
R:Prost, E.; Deryckere, F.; Roos, C.; Haenlin, M.; Pantescio, V.; Mohler, E.
A:Title: Role of the oocyte nucleus in determination of the dorsoventral polarity of Dros
A:Reference number: A28826
A:Accession: A28826
A:Molecule type: mRNA
A:Residues: 1-216, 'HH', 279-281, 'VDHHR', 287-463 <PRO>
A:Cross-references: GB:X12836; NID:g8148; PID:g295771
C:Genetics:
A:Gene: FlyBase:fs(1)K10
A:Cross-references: FlyBase:FBgn0000810
A:Map position: X
A:Introns: 432/3
C:Keywords: DNA binding

Query Match      42.7%; Score 53; DB 2; Length 463;
Best Local Similarity 47.8%; Pred. No. 11;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY          1 YEKKNAYQKANOAVLKAKEASS 23
||||| ||||| :|: |||||
Db          414 YEAAAEVYLKQEAATVKAKDAKS 436
||||| ||||| :|: |||||

RESULT 7
T22984
hypothetical protein F59B8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22984
R:Matthews, P.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19647
A:Accession: T22984
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-213 <WIL>
A:Cross-references: EMBL:Z68343; PIDN:CAA92779.1; GSPDB:GN00022; CESP:F59B8.1
A:Experimental source: clone F59B8
C:Genetics:
A:Gene: CESP:F59B8.1
```

A;Map position: 4
A;Introns: 31/2; 65/3; 94/2; 138/1; 170/3

Query Match 41.9%; Score 52; DB 2; Length 213;
Best Local Similarity 43.5%; Pred. No. 6.8;
Matches 10; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKKASS 23
| : ||| : : ||| : : ||| : :
Db 30 YRRLKNALKKQEDYLSKEANA 52

RESULT 8
T48234
hypothetical protein T7H20.180 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 05-May-2000
C;Accession: T48234
R;Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.; Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24488
A;Accession: T48234
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-420 <BEV>
A;Cross-references: EMBL:AL162508
A;Experimental source: cultivar Columbia; BAC clone T7H20
C;Genetics:
A;Map position: 5
A;Introns: 67/1; 109/2; 134/1; 149/3; 196/1; 211/1; 274/2; 295/1; 328/3; 347/1; 362/1
A;Note: T7H20.180
C;Superfamily: Arabidopsis thaliana hypothetical protein T7H20.180

Query Match 41.9%; Score 52; DB 2; Length 420;
Best Local Similarity 52.4%; Pred. No. 13;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKKA 21
| : ||| : : ||| : : ||| : :
Db 227 FELAKEMYQKAIQGVTTKES 247

RESULT 9
B82363
gIpe protein VC0100 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82363
R;Maidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
L. R.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: B82363
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-106 <HEI>
A;Cross-references: GB:AE004101; GB:AE003852; NID:g9654497; PIDN:AAF93278.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0100
A;Map position: 1
C;Superfamily: gIpe protein

Query Match 40.3%; Score 50; DB 2; Length 106;
Best Local Similarity 40.9%; Pred. No. 6.4;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 4 AKNAYQKANOAVLKAKKASSYD 25
| : ||| : : ||| : : ||| : :
Db 37 AKNAYHLTNSQVQFMEQAEFD 58

RESULT 10

B85432
hypothetical protein AT4g36600 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: B85432
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: B85432
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-347 <STO>
A;Cross-references: GB:NC_001368; NID:g7270608; PIDN:CAB80326.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g36600
A;Map position: 4

Query Match 39.9%; Score 49.5; DB 2; Length 347;
Best Local Similarity 36.8%; Pred. No. 24;
Matches 14; Conservative 4; Mismatches 7; Indels 13; Gaps 1;

Qy 1 YEKAKNAYQKANOAVLKKA-----KEASSYD 25
| : ||| : : ||| : : ||| : :
Db 145 YDKAGSAYEKAGQAKDMAYDKAGQAKDMAYDKVGSAYD 182

RESULT 11

F97149
probable flavodoxin [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: F97149
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: F97149
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-180 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK79985.1; PID:g15025010; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC2026

Query Match 39.5%; Score 49; DB 2; Length 180;
Best Local Similarity 58.8%; Pred. No. 15;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 6 NAYQKANOAVLKAKKAS 22
| : ||| : : ||| : : ||| : :
Db 28 NKPRQANQKILKAKEAN 44

RESULT 12

T15624
hypothetical protein C25H3.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15624
R;Johnson, D.

submitted to the EMBL Data Library, June 1995
A;Description: The sequence of C. elegans cosmid C25H3.
A;Reference number: Z18379
A;Accession: T15624
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-651 <JOH>
A;Cross-references: EMBL:U29535; NID:g868251; PID:g868256; PIDN:AAA68785.1; CESP:C25H3.6

A;Experimental source: strain Bristol N2

C;Genetics:

A;Gene: CESP:C25H3.6

A;Introns: 58/1; 213/3; 273/1; 446/2; 468/2; 484/2; 510/1; 552/1; 609/1

Query Match 39.5%; Score 49; DB 2; Length 651;

Best Local Similarity 40.0%; Pred. No. 53;

Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 6 NAYQKAOVLKAKEASSYD 25

|::| |::| |::|

Db 526 NSQSFNMAVAPROAPDYD 545

RESULT 13

A24341

C;Species: Pseudomonas aeruginosa transposon Tn501

C;Genetics: Pseudomonas aeruginosa

A;Date: 29-Aug-1987 #sequence_revision 29-Aug-1987 #text_change 21-Jul-2000

C;Accession: A24341

R;Brown, N.L.; Winnie, J.N.; Fritzinger, D.; Fridmore, R.D.

Nucleic Acids Res. 13, 5657-5669, 1985

A;Title: The nucleotide sequence of the trpA gene completes the sequence of the Pseudom

A;Reference number: A24341; MUID:85297764; PMID:2994007

A;Accession: A24341

A;Molecule type: DNA

A;Residues: 1-988 <BRO>

A;Cross-references: GB:Z00027; GB:K00031; GB:K01725; GB:X01297; GB:X03406; NID:g43714; F

C;Genetics:

A;Gene: trpA

C;Superfamily: transposase Tn21

C;Keywords: DNA binding

Query Match 38.7%; Score 48; DB 2; Length 988;

Best Local Similarity 62.5%; Pred. No. 11e+02;

Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 YOKAQOAVLKAKEASS 23

|::| |::| |::|

Db 330 YSKIGQALLKAKEAGS 345

RESULT 14

R5EC23

C;Species: Escherichia coli

A;Date: 28-Feb-1980 #sequence_revision 31-Oct-1997 #text_change 01-Mar-2002

C;Accession: A65125; D23129; A02814

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: A65125

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-100 <BLAT>

A;Cross-references: GB:AE000408; GB:U00096; NID:g1789694; PIDN:AACT6343.1; PID:g1789714;

A;Experimental source: strain K-12, substrain MGL655

R;Zurawski, G.; Zurawski, S.M.

Nucleic Acids Res. 13, 4521-4526, 1985

A;Title: Structure of the Escherichia coli S10 ribosomal protein operon.

A;Reference number: A23129; MUID:85242118; PMID:3892488

A;Accession: D23129

A;Molecule type: DNA

A;Residues: 1-100 <ZUR>

A;Cross-references: GB:X02613; NID:g42825; PIDN:CAA26462.1; PID:g534975

A;Note: translation of initiator Met is not shown; the initiator overlaps the last codon

R;Wittmann-Liebold, B.; Greuer, B.

FEBS Lett. 108, 69-74, 1979

A;Title: Primary structure of protein L23 from the Escherichia coli ribosome.

A;Reference number: A02814; MUID:80092111; PMID:391594

A;Accession: A02814

A;Molecule type: protein

A;Residues: 1-79, 81-100 <WIT>

A;Experimental source: strain K12

R;Arnold, R.J.; Reilly, J.P.

Anal. Biochem. 269, 105-112, 1999

A;Title: Observation of Escherichia coli ribosomal proteins and their posttranslational n

A;Reference number: A59071; MUID:99196679; PMID:10094780

A;Contents: annotation; mass spectrographic analysis

A;Note: no post-translational modifications were observed in mass spectrographic analysis

C;Genetics:

A;Gene: rplW

A;Map position: 73 min

C;Complex: the ribosome is composed of the large (50S) and small (30S) subunit; the large

S rRNA and 22 distinct proteins

C;Complex: large subunit ribosomal proteins: L1 (PIR:R5EC1), L3 (PIR:R5EC3), L2 (PIR:R5EC

(PIR:R5EC11), L13 (PIR:R5EC13), L14 (PIR:R5EC14), L15 (PIR:R5EC15), L16 (PIR:R5EC16), L1;

C;Function:

A;Pathway: protein biosynthesis

C;Superfamily: Escherichia coli ribosomal protein L23

C;Keywords: protein biosynthesis; ribosome

F;1-100/Product: ribosomal protein L23 #status experimental <MAT>

Query Match 38.3%; Score 47.5; DB 1; Length 100;

Best Local Similarity 54.5%; Pred. No. 13;

Matches 12; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 2 EKAKYQKAOAVLK-AKEAS 22

|::| |::| |::|

Db 18 EKASTAMEKSNTIVLKVANDAT 39

RESULT 15

G91151

S0S ribosomal subunit protein L23 [imported] - Escherichia coli (strain O157:H7, substr

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C;Accession: G91151

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;

Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genom

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: G91151

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-100 <HAY>

A;Cross-references: GB:BA000007; PIDN:BA837606.1; PID:g13363656; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs4183

C;Superfamily: Escherichia coli ribosomal protein L23

Query Match 38.3%; Score 47.5; DB 2; Length 100;

Best Local Similarity 54.5%; Pred. No. 13;

Matches 12; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

QY 2 EKAKYQKAOAVLK-AKEAS 22

|::| |::| |::|

Db 18 EKASTAMEKSNTIVLKVANDAT 39

Search completed: September 22, 2004, 18:09:58

Job time : 13.5041 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 11.7769 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-11

Perfect score: 124

Sequence: 1 YERAKNAYQKANOAVLKAKEASSYD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	93.5	64	3	US-08-416-711-1
2	116	93.5	64	4	US-09-356-497-1
3	96	77.4	23	3	US-08-416-711-2
4	96	77.4	23	4	US-09-356-497-2
5	50	40.3	28	1	US-08-303-025-12
6	50	40.3	28	2	US-08-436-703B-1
7	50	40.3	29	1	US-08-152-488-10
8	50	40.3	29	1	US-08-152-488-11
9	50	40.3	29	1	US-08-303-025-10
10	50	40.3	29	1	US-08-303-025-11
11	50	40.3	29	1	US-08-303-025-13
12	50	40.3	29	1	US-08-677-304-11
13	50	40.3	29	1	US-08-677-304-11
14	50	40.3	29	2	US-08-436-703B-3
15	50	40.3	29	2	US-08-436-703B-15
16	50	40.3	32	1	US-08-152-488-13
17	50	40.3	32	1	US-08-303-025-15
18	50	40.3	32	1	US-08-677-304-13
19	50	40.3	32	2	US-08-436-703B-2
20	50	40.3	33	1	US-08-303-025-16
21	50	40.3	33	2	US-08-436-703B-4
22	49	39.5	29	1	US-08-152-488-12
23	49	39.5	29	1	US-08-303-025-14
24	49	39.5	29	1	US-08-677-304-12
25	49	39.5	29	2	US-08-436-703B-16
26	49	39.5	928	4	US-09-134-000C-6590
27	47.5	38.3	100	4	US-09-732-210-745

Sequence 308, Appl
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 11042, A
Sequence 36, Appl
Sequence 7464, Ap
Sequence 8019, Ap
Sequence 32, Appl
Sequence 7010, Ap
Sequence 4752, Ap
Sequence 7225, Ap
Sequence 8679, Ap
Sequence 3699, Ap
Sequence 4511, Ap
Sequence 9, Appli
Sequence 3847, Ap

ALIGNMENTS

RESULT.1
US-08-416-711-1
; Sequence 1, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P. C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-416-711-1

Query Match 93.5%; Score 116; DB 3; Length 64;
Best Local Similarity 92.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 2

US-09-356-497-1
; Sequence 1, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 93.5%; Score 116; DB 4; Length 64;
Best Local Similarity 92.0%; Pred. No. 4.3e-10;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 3

US-08-416-711-2
; Sequence 2, Application US/08416711

; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
; US-08-416-711-2

Query Match 77.4%; Score 96; DB 3; Length 23;
Best Local Similarity 84.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYOKANQAVL--KEASSYD 23

RESULT 4

US-09-356-497-2
; Sequence 2, Application US/09356497
; Patent No. 6472519
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE

; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-356-497-2

Query Match 77.4%; Score 96; DB 4; Length 23;
Best Local Similarity 84.0%; Pred. No. 1.2e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 1;
QY 1 YEKAKYQKQANQAVLKAKESASYD 25
Db 1 HERAKYQKQANQAVL--KESASYD 23

RESULT 5
US-08-303-025-12
; Sequence 12, Application US/08303025
; Patent No. 5614494
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 150 West Jefferson, Suite 2500
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226-4415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS v.6.22
; SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,025
; FILING DATE: 08-SEPT-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/06829
; FILING DATE: 14-AUG-1992
; APPLICATION NUMBER: US 08/152,488
; FILING DATE: 12-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REFERENCE/DOCKET NUMBER: 7WK-060548-00231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-496-7622
; TELEFAX: 313-496-8454
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
US-08-303-025-12

Query Match 40.3%; Score 50; DB 1; Length 28;
Best Local Similarity 52.4%; Pred. No. 0.74;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 2 EKAKYQKQANQAVLKAKESAS 22
Db 3 KKAKKAQKAKKAAKKAKKAA 23

RESULT 6
US-08-436-703B-1
; Sequence 1, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; STREET: Suite 1525
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6;
; SOFTWARE: ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; US-08-436-703B-1

Query Match 40.3%; Score 50; DB 2; Length 28;
Best Local Similarity 52.4%; Pred. No. 0.74;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKEAS 22
Db 3 KKAKKAARKAKKAAKAKKAA 23

RESULT 7
US-08-152-488-10
; Sequence 10, Application US/08152488
; Patent No. 5534619
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6; ASCII (DOS) Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,488
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7WG
; TELEPHONE: 908-276-5543
; TELEFAX: 908-276-5344
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKEAS 22
Db 7 KKAKKAARKAKKAAKAKKAA 27

RESULT 8
US-08-152-488-11
; Sequence 11, Application US/08152488
; Patent No. 5534619
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 512 Springfield Avenue
; CITY: Cranford
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07016-1811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6; ASCII (DOS) Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,488
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: RM-7WG
; TELEPHONE: 908-276-3344
; TELEFAX: 908-276-5543
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: N/A
; PUBLICATION INFORMATION:
; AUTHORS: N/A
; TITLE: N/A
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: PCT/US92/08069
; FILING DATE: 14-AUG-1993
; US-08-152-488-11

Query Match 40.3%; Score 50; DB 1; Length 29;
```

; AFFILIANT: STANLEY, JAMES C.
 ; TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
 ; TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN

;/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL
;/ NUMBER OF SEQUENCES: 16
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Benita J. Rohm, Esq.
;/ STREET: 150 West Jefferson, Suite 2500
;/ CITY: Detroit
;/ STATE: Michigan
;/ COUNTRY: United States of America
;/ ZIP: 48226-4415
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy diskette 3.5" 1.44Mb
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: MS-DOS v.6.22
;/ SOFTWARE: WordPerfect 6.1; ASCII (DOS)Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/303,025
;/ FILING DATE: 08-SEPT-1994
;/ CLASSIFICATION: 514
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: PCT/US92/06829
;/ FILING DATE: 14-AUG-1992
;/ APPLICATION NUMBER: US 08/152,488
;/ FILING DATE: 12-NOV-1993
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Rohm, Benita J.
;/ REFERENCE/DOCKET NUMBER: 7WH-060548-00231
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 313-496-7622
;/ TELEFAX: 313-496-8454
;/ INFORMATION FOR SEQ ID NO: 13:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 29 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: N/A
;/ TOPOLOGY: N/A
;/ MOLECULE TYPE: peptide
;/ ORIGINAL SOURCE:
;/ ORGANISM: N/A
;/ PUBLICATION INFORMATION:
;/ AUTHORS: N/A
;/ TITLE: N/A
;/ DOCUMENT NUMBER: PCT/US92/08069
;/ FILING DATE: 14-AUG-1993
;/ US-08-303-025-13

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAEAS 22
DB 7 KKAKKAQKAKKAKKAKKAA 27

RESULT 12
US-08-677-304-10
;/ Sequence 10, Application US/08677304
;/ Patent No. 5721212
;/ GENERAL INFORMATION:
;/ APPLICANT: Wakefield, Thomas W.
;/ APPLICANT: Andrews, Philip C.
;/ APPLICANT: Stanley, James C.
;/ TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
;/ TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
;/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL
;/ NUMBER OF SEQUENCES: 13
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Benita J. Rohm, Esq.
;/ STREET: 512 Springfield Avenue
;/ CITY: Cranford
;/ STATE: New Jersey
;/ COUNTRY: United States of America
;/ ZIP: 07016-1811

;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: MS-DOS
;/ SOFTWARE: WordPerfect 6; ASCII (DOS)Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/677,304
;/ FILING DATE:
;/ CLASSIFICATION: 530
;/ PRIOR APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/152,488
;/ FILING DATE: 12-NOV-1993
;/ APPLICATION NUMBER: PCT/US92/08069
;/ FILING DATE: 14-AUG-1993
;/ ATTORNEY/AGENT INFORMATION:
;/ NAME: Rohm, Benita J.
;/ REGISTRATION NUMBER: 28,664
;/ REFERENCE/DOCKET NUMBER: RM-7WG
;/ TELECOMMUNICATION INFORMATION:
;/ TELEPHONE: 908-276-3344
;/ TELEFAX: 908-276-5543
;/ INFORMATION FOR SEQ ID NO: 10:
;/ SEQUENCE CHARACTERISTICS:
;/ LENGTH: 29 amino acids
;/ TYPE: amino acid
;/ STRANDEDNESS: No. 5721212 Relevant
;/ TOPOLOGY: No. 5721212 Relevant
;/ MOLECULE TYPE: peptide
;/ ORIGINAL SOURCE:
;/ ORGANISM: N/A
;/ PUBLICATION INFORMATION:
;/ AUTHORS: N/A
;/ TITLE: N/A
;/ PUBLICATION INFORMATION:
;/ DOCUMENT NUMBER: PCT/US92/08069
;/ FILING DATE: 14-AUG-1993
;/ US-08-677-304-10

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAEAS 22
DB 7 KKAKKAQKAKKAKKAKKAA 27

RESULT 13
US-08-677-304-11
;/ Sequence 11, Application US/08677304
;/ Patent No. 5721212
;/ GENERAL INFORMATION:
;/ APPLICANT: Wakefield, Thomas W.
;/ APPLICANT: Andrews, Philip C.
;/ APPLICANT: Stanley, James C.
;/ TITLE OF INVENTION: NOVEL PEPTIDES FOR HEPARIN AND
;/ TITLE OF INVENTION: LOW MOLECULAR WEIGHT HEPARIN
;/ TITLE OF INVENTION: ANTICOAGULATION REVERSAL
;/ NUMBER OF SEQUENCES: 13
;/ CORRESPONDENCE ADDRESS:
;/ ADDRESSEE: Benita J. Rohm, Esq.
;/ STREET: 512 Springfield Avenue
;/ CITY: Cranford
;/ STATE: New Jersey
;/ COUNTRY: United States of America
;/ ZIP: 07016-1811
;/ COMPUTER READABLE FORM:
;/ MEDIUM TYPE: Floppy disk
;/ COMPUTER: IBM PC compatible
;/ OPERATING SYSTEM: MS-DOS
;/ SOFTWARE: WordPerfect 6; ASCII (DOS)Text
;/ CURRENT APPLICATION DATA:
;/ APPLICATION NUMBER: US/08/677,304

;; FILING DATE: 530
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/152,488
;; FILING DATE: 12-NOV-1993
;; APPLICATION NUMBER: PCT/US92/08069
;; FILING DATE: 14-AUG-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rohm, Benita J.
;; REGISTRATION NUMBER: 28,664
;; REFERENCE/DOCKET NUMBER: RM-7WG
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908-276-3344
;; TELEFAX: 908-276-5543
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: No. 5721212 Relevant
;; TOPOLOGY: No. 5721212 Relevant
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: N/A
;; PUBLICATION INFORMATION:
;; AUTHORS: N/A
;; TITLE: N/A
;; PUBLICATION INFORMATION:
;; DOCUMENT NUMBER: PCT/US92/08069
;; FILING DATE: 14-AUG-1993
US-08-677-304-11

Query Match 40.3%; Score 50; DB 1; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKKAS 22
Db 7 KKAKKAACKAKKAACKAKAA 27

RESULT 14
US-08-436-703B-3
; Sequence 3, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6;
; SOFTWARE: ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A

;; FILING DATE: N/A
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rohm, Benita J.
;; REGISTRATION NUMBER: 28,664
;; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 313-965-1976
;; TELEFAX: 313-965-1951
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 29 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: N/A
;; TOPOLOGY: N/A
;; MOLECULE TYPE: peptide
;; ORIGINAL SOURCE:
;; ORGANISM: N/A
;; PUBLICATION INFORMATION:
;; AUTHORS: N/A
;; TITLE: N/A
US-08-436-703B-3

Query Match 40.3%; Score 50; DB 2; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKKAS 22
Db 7 KKAKKAACKAKKAACKAKAA 27

RESULT 15
US-08-436-703B-15
; Sequence 15, Application US/08436703B
; Patent No. 5919761
; GENERAL INFORMATION:
; APPLICANT: Wakefield, Thomas W.
; APPLICANT: Andrews, Philip C.
; APPLICANT: Stanley, James C.
; TITLE OF INVENTION: NOVEL PEPTIDES FOR
; TITLE OF INVENTION: HEPARIN AND LOW MOLECULAR
; TITLE OF INVENTION: WEIGHT HEPARIN
; TITLE OF INVENTION: ANTICOAGULATION REVERSAL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benita J. Rohm, Esq.
; STREET: 6601 Woodward Avenue
; CITY: Detroit
; STATE: Michigan
; COUNTRY: United States of America
; ZIP: 48226
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk 1.44Mb, 3.5"
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 6;
; SOFTWARE: ASCII (DOS)Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,703B
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: N/A
; FILING DATE: N/A
; ATTORNEY/AGENT INFORMATION:
; NAME: Rohm, Benita J.
; REGISTRATION NUMBER: 28,664
; REFERENCE/DOCKET NUMBER: 7WK-060548-00233
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313-965-1976
; TELEFAX: 313-965-1951
; INFORMATION FOR SEQ ID NO: 15:

```

SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: N/A
PUBLICATION INFORMATION:
AUTHOR: N/A
TITLE: N/A
US-08-436-7038-15

```

Query Match 40.3%; Score 50; DB 2; Length 29;
Best Local Similarity 52.4%; Pred. No. 0.77;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy [REDACTED] 2 EKAKYAYQKANQAVLKAKEAS 22
[REDACTED] :||| | :||| : ||| :
D [REDACTED] 7 KKAUKAAKKAUKAUKAKKAAA 27

Search completed: September 22, 2004, 18:11:00
Job time : 14.7769 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	124	100.0	25	14	US-10-294-770-11	Sequence 11, Appl	
2	124	100.0	25	16	US-10-774-602-11	Sequence 11, Appl	
3	116	93.5	64	14	US-10-294-770-1	Sequence 1, Appl	
4	116	93.5	64	16	US-10-238-741-1	Sequence 1, Appl	
5	116	93.5	64	16	US-10-774-602-1	Sequence 1, Appl	
6	96	77.4	23	14	US-10-294-770-2	Sequence 2, Appl	
7	96	77.4	23	16	US-10-238-741-2	Sequence 2, Appl	
8	96	77.4	23	16	US-10-774-602-2	Sequence 2, Appl	
9	51.5	41.5	866	16	US-10-437-963-203902	Sequence 203902,	
10	51.5	41.5	1109	16	US-10-437-963-203905	Sequence 203905,	
11	50	40.3	79	12	US-10-393-449-16	Sequence 16, Appl	
12	50	40.3	79	12	US-10-393-449-20	Sequence 20, Appl	
13	50	40.3	79	12	US-10-393-449-66	Sequence 66, Appl	
14	50	40.3	79	12	US-10-393-449-70	Sequence 70, Appl	
15	50	40.3	79	14	US-10-177-725-16	Sequence 16, Appl	

```

; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-774-602-11

Query Match          100.0%; Score 124; DB 16; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 YEKAKNAYOKANQAVLKAKEASSYD 25

RESULT 3
US-10-294-770-1
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 64
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-1

Query Match          93.5%; Score 116; DB 14; Length 64;
Best Local Similarity 92.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 4
US-10-238-741-1
; Publication No. US10238741
; Sequence 1, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE

```

```

; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238,741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-238-741-1

Query Match          93.5%; Score 116; DB 16; Length 64;
Best Local Similarity 92.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 5
US-10-774-602-1
; Sequence 1, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08

```


;; PRIOR APPLICATION NUMBER: PCT/FR93/01024
;; PRIOR FILING DATE: 1993-10-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 1
;; LENGTH: 64
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-10-774-602-1

Query Match 93.5%; Score 116; DB 16; Length 64;
Best Local Similarity 92.0%; Pred. No. 2.3e-09;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYQKANOAVLKAKEASSYD 25

RESULT 6
US-10-294-770-2
Sequence 2, Application US/10294770
Publication No. US20030161840A1
GENERAL INFORMATION:
;; APPLICANT: DRUILHE, PIERRE
;; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
;; FILE REFERENCE: 230759USOCIP
;; CURRENT APPLICATION NUMBER: US/10/294,770
;; CURRENT FILING DATE: 2002-11-15
;; PRIOR APPLICATION NUMBER: US 09/356,947
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: US 08/416,711
;; PRIOR FILING DATE: 1995-08-08
;; PRIOR APPLICATION NUMBER: PCT/FR93/01024
;; PRIOR FILING DATE: 1993-10-18
;; NUMBER OF SEQ ID NOS: 14
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 2
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-10-294-770-2

Query Match 77.4%; Score 96; DB 14; Length 23;
Best Local Similarity 84.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 YEKAKNAYQKANOAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYQKANOAVLKAKEASSYD 23

RESULT 7
US-10-238-741-2
Sequence 2, Application US/10238741
Publication No. US20040096466A1
GENERAL INFORMATION:
;; APPLICANT: DRUILHE, PIERRE
;; BOUHAROUN-TAYOUN, HASNAQ
;; OEUVRAY, CLAUDE
;; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
;; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;; CITY: ARLINGTON
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22202
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/10/238,741
;; FILING DATE: 09-Nov-2002
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/356,497
;; FILING DATE: 19-Jul-1999
;; APPLICATION NUMBER: US/08/416,711
;; FILING DATE: 08-AUG-1995
;; APPLICATION NUMBER: PCT/FR93/01024
;; FILING DATE: 18-OCT-1993
;; APPLICATION NUMBER: FR 92/12488
;; FILING DATE: 19-OCT-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NORMAN F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220

;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 23 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-238-741-2

Query Match 77.4%; Score 96; DB 16; Length 23;
Best Local Similarity 84.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 YEKAKNAYQKANOAVLKAKEASSYD 25
:|:|||||
Db 1 HERAKNAYQKANOAVLKAKEASSYD 23

RESULT 8
US-10-774-602-2
Sequence 2, Application US/10774602
Publication No. US20040141987A1
GENERAL INFORMATION:

;; APPLICANT: DRUILHE, PIERRE
;; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
;; FILE REFERENCE: 248791USODIV
;; CURRENT APPLICATION NUMBER: US/10/774,602
;; CURRENT FILING DATE: 2004-02-10
;; PRIOR APPLICATION NUMBER: US 09/356,947
;; PRIOR FILING DATE: 1999-07-19
;; PRIOR APPLICATION NUMBER: US 10/238,741
;; PRIOR FILING DATE: 2002-09-11
;; PRIOR APPLICATION NUMBER: US 08/416,711
;; PRIOR FILING DATE: 1995-08-08
;; PRIOR APPLICATION NUMBER: PCT/FR93/01024
;; PRIOR FILING DATE: 1993-10-18
;; SOFTWARE: PatentIn version 3.1
;; NUMBER OF SEQ ID NOS: 14
;; SEQ ID NO 2
;; LENGTH: 23
;; TYPE: PRT
;; ORGANISM: Plasmodium falciparum
US-10-774-602-2

Query Match 77.4%; Score 96; DB 16; Length 23;
Best Local Similarity 84.0%; Pred. No. 5.5e-07;
Matches 21; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

Qy 1 YEKAKNAYQKANOAVLKAKEASSYD 25


```
Best Local Similarity 63.2%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 KAKNAYQKQANQAVLKAKEA 21
   ||| | | | | | | | | |
Db 58 KAKEAEAKAKEAEAKAKEA 76

RESULT 13
US-10-393-449-66
; Sequence 66, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: RIGL-007CIP3
; CURRENT APPLICATION NUMBER: US/10/177,725
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/177,725
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 66
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; NAME/KEY: MISC FEATURE
; LOCATION: (21)..(52)
; OTHER INFORMATION: "Xaa" at positions 21-23, 25-27, 29-30, 32-34, 36-37, 39-41, 43-4
; OTHER INFORMATION: 5, 47-48, and 50-52 can be any amino acid
US-10-393-449-66

Query Match 40.3%; Score 50; DB 12; Length 79;
Best Local Similarity 63.2%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 KAKNAYQKQANQAVLKAKEA 21
   ||| | | | | | | | | |
Db 58 KAKEAEAKAKEAEAKAKEA 76

RESULT 14
US-10-393-449-70
; Sequence 70, Application US/10393449
; Publication No. US20030224412A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: RIGL-007CIP3
; CURRENT APPLICATION NUMBER: US/10/393,449
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/177,725
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 70
; LENGTH: 79
; TYPE: PRT
```

```
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; NAME/KEY: MISC FEATURE
; LOCATION: (21)..(52)
; OTHER INFORMATION: "Xaa" at positions 21-23, 25-27, 29-30, 32-34, 36-37, 39-41, 43-4
; OTHER INFORMATION: 5, 47-48 and 50-52 can be any amino acid
US-10-393-449-70

Query Match 40.3%; Score 50; DB 12; Length 79;
Best Local Similarity 63.2%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 KAKNAYQKQANQAVLKAKEA 21
   ||| | | | | | | | | |
Db 58 KAKEAEAKAKEAEAKAKEA 76

RESULT 15
US-10-177-725-16
; Sequence 16, Application US/10177725
; Publication No. US20030143562A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S
; FILE REFERENCE: A-66900-4/RMS/AMS
; CURRENT APPLICATION NUMBER: US/10/177,725
; CURRENT FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 09/415,765
; PRIOR FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: US 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 16
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-177-725-16

Query Match 40.3%; Score 50; DB 14; Length 79;
Best Local Similarity 63.2%; Pred. No. 9.6;
Matches 12; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 KAKNAYQKQANQAVLKAKEA 21
   ||| | | | | | | | | |
Db 58 KAKEAEAKAKEAEAKAKEA 76
```

Search completed: September 22, 2004, 18:34:03
Job time : 62.9256 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:02:42 ; Search time 10.6446 Seconds
(without alignments)
253.026 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147
Sequence: 1 PEHKEENMLSHLYVSSKDKENISKEND 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:
1: pir1:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.5	37.1	1155	2 B96761	probable protein k
2	54	36.7	839	2 T12827	hypothetical prote
3	53	36.1	351	2 T17246	hypothetical prote
4	52.5	35.7	558	2 T18467	hypothetical prote
5	52	35.4	749	2 A57276	MAD1 protein - yea
6	52	35.4	993	2 S49461	synaptonemal compl
7	52	35.4	1249	2 H71404	hypothetical prote
8	52	35.4	1475	2 F86399	protein F17L21.22
9	52	35.4	2206	2 G71611	hypothetical prote
10	51.5	35.0	407	2 A71613	hypothetical prote
11	51	34.7	492	2 S50554	hypothetical prote
12	51	34.7	572	1 DEMSNX	malate dehydrogena
13	50.5	34.4	1748	2 S42136	cnjB protein - Tet
14	50.5	34.4	3724	2 T18427	hypothetical prote
15	50	34.0	541	1 S45428	PF112 protein - y
16	50	34.0	1157	2 C96761	hypothetical prote
17	50	34.0	1225	2 T18954	hypothetical prote
18	50	34.0	1997	2 F71607	DNA helicase II BR
19	49.5	33.7	265	2 AF3381	NAD kinase (EC 2.7
20	49.5	33.7	1714	2 E71609	Ser/Thr protein ki
21	49.5	33.7	3394	2 T18501	hypothetical prote
22	49	33.3	646	2 F71620	hypothetical prote
23	49	33.3	757	1 T50959	probable endopolyp
24	49	33.3	946	2 S28061	SCP1 protein - rat
25	49	33.3	1817	2 H71611	probable secreted
26	49	33.3	1984	2 A44396	p-type cation tran
27	48.5	33.0	769	2 F89870	serine proteinase
28	48.5	33.0	1939	2 T18372	repeat organellar
29	48	32.7	417	2 E75148	hypothetical prote

30	48	32.7	495	2 E70198	hypothetical prote
31	48	32.7	845	2 I48176	synaptonemal compl
32	48	32.7	1075	2 T00341	hypothetical prote
33	48	32.7	1398	2 H71606	hypothetical prote
34	48	32.7	2295	2 B71621	probable membrane
35	47.5	32.3	370	2 S57716	RNA dependent RNA
36	47.5	32.3	378	2 T18486	hypothetical prote
37	47.5	32.3	643	2 S35511	high mobility grou
38	47.5	32.3	889	2 A41259	potassium transpor
39	47.5	32.3	1017	2 T18488	hypothetical prote
40	47.5	32.3	1073	2 S14032	kinesin-related pr
41	47.5	32.3	1085	2 T38378	kinesin-like prote
42	47	32.0	204	2 S56256	hypothetical prote
43	47	32.0	399	2 T38388	hypothetical wd-40
44	47	32.0	482	2 B64317	hypothetical prote
45	47	32.0	508	2 S74977	hypothetical prote

ALIGNMENTS

RESULT 1

B96761

Probable protein kinase T9L24.36 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: B96761

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.B.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: B96761

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1155 <STO>

A;Cross-references: GB:AE005173; NID:gl1120796; PIDN:AAG30976.1; GSPDB:GN00141

C;Genetics:

A;Gene: T9L24.36

A;Map position: 1

Query Match 37.1%; Score 54.5; DB 2; Length 1155;

Best Local Similarity 48.1%; Pred. No. 38;

Matches 13; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 3 HKKEENMLS-HLYVSSKDKENISKEND 28

Db 472 YAEDESYLSGERYLQSKDAEPISSEND 498

RESULT 2

T12827

Hypothetical protein yonO - Bacillus subtilis phage SPBc2

C;Species: Bacillus subtilis phage SPBc2

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 20-Jun-2000

C;Accession: T12827; G69914

R;Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mael, C.; Karamata, D.

submitted to the EMBL Data Library, August 1997

A;Description: The complete nucleotide sequence of the Bacillus subtilis SPBc2 prophage

A;Reference number: Z17583

A;Accession: T12827

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-839 <IAZ>

A;Cross-references: EMBL:AF020713; NID:G3025478; PID:G3025541; PIDN:AAC13036.1

R;Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
keuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Toseato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: G69914
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-839 <KUN>
A:Cross-references: GB:Z99115; GB:AL009126; NID:G2634478; PIDN:CAB14022.1; PID:G2634524
A:Experimental source: strain 168
C:Genetics:
C:Superfamily: *Bacillus subtilis* phase SPBc2 hypothetical protein yonO
Query Match 36.7%; Score 54; DB 2; Length 839;
Best Local Similarity 33.3%; Pred. No. 32;
Matches 11; Conservative 8; Mismatches 8; Indels 6; Gaps 1;
QY 1 PEHKKE-----ENMLSHLYVSSKDKENISKEN 27
DB 636 PDHVEKIDKYNKVPHPFINAKDKEHSVES 668
RESULT 3
T17246
hypothetical protein DKFZp586M0617.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
C:Accession: T17246
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
A:Accession: T17246
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-351 <KOE>
A:Experimental source: adult uterus; clone DKFZp586M0617
C:Genetics:
A:Note: DKFZp586M0617.1
C:Superfamily: *Caenorhabditis elegans* hypothetical protein F30A10.3
Query Match 36.1%; Score 53; DB 2; Length 351;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 13; Conservative 5; Mismatches 7; Indels 14; Gaps 1;
QY 2 EHKKEENMLSH-----LYVSSKDKENISKE 26
DB 59 QHRKEEMKSHLKEEFWELKKSEVLTYTVKKGNISSQ 97
RESULT 4
T18467
hypothetical protein C0465c - malaria parasite (*Plasmodium falciparum*)
C:Species: *Plasmodium falciparum*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2002
C:Accession: T18467
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z18937
A:Accession: T18467
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-558 <LAW>
A:Cross-references: EMBL:AL008970; NID:e1407852; PIDN:CAA15610.1
C:Genetics:
A:Map position: 3
A:Introns: 84/1; 160/1
A:Note: C0465c
Query Match 35.7%; Score 52.5; DB 2; Length 558;
Best Local Similarity 34.5%; Pred. No. 33;
Matches 10; Conservative 10; Mismatches 4; Indels 5; Gaps 1;
QY 5 KEENMLSHLY-----VSSKDKENISKEND 28
DB 149 ENENINENVYENKDISNKDKEHVSHQNE 177
RESULT 5
A57276
MAD1 protein - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: protein G3191; protein YGL086w
C:Species: *Saccharomyces cerevisiae*
C:Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 21-Jul-2000
C:Accession: A57276; S64093; S48246
R:Hardwick, K.G.; Murray, A.W.
J. Cell Biol. 131, 709-720, 1995
A:Title: Mad1p, a phosphoprotein component of the spindle assembly checkpoint in budding
A:Reference number: A57276; MUID:96042315; PMID:7593191
A:Accession: A57276
A:Molecule type: DNA
A:Residues: 1-749 <HAR>
A:Cross-references: GB:U14632; NID:g551091; PIDN:AAA91620.1; PID:g551092
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64093
A:Molecule type: DNA
A:Residues: 1-749 <RIB>
A:Cross-references: EMBL:Z72608; NID:g1322610; PID:e243297; PID:g1322611; MIPS:YGL086w
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:MAD1
A:Cross-references: SGD:S0003054; MIPS:YGL086w
A:Map position: 7L
C:Function:
A:Description: component of spindle assembly checkpoint which prevents cells from initial
C:Keywords: cell division control; nucleus; phosphoprotein
Query Match 35.4%; Score 52; DB 2; Length 749;
Best Local Similarity 42.3%; Pred. No. 52;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 2 EHKKEENMLSHLYVSSKDKENISKEN 27
DB 283 ENEKLQNKLSQLHVESQVENLQLEN 308
RESULT 6
S49461
synaptonemal complex protein 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: S49461; S59599
R:Julien, S.; Luc, M.; Francois, C.
submitted to the EMBL Data Library, October 1994
A:Description: Cloning and sequencing of the murine SCP1 cDNA.
A:Reference number: S49461
A:Accession: S49461
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-993 <JUL>
A:Cross-references: EMBL:Z38118; NID:g1360015; PIDN:CAA86262.1; PID:g558603
R:Sage, J.; Martin, L.; Cuzin, F.; Rassoulzadegan, M.
Biochim. Biophys. Acta 1263, 258-260, 1995

hypothetical protein YER051w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 19-Apr-2002
C:Accession: S50554
R:Dieterich, F.S.
A:Description: The sequence of S. cerevisiae cosmid 9379, 9581, and lambda clone 4678.
A:Reference number: S50432
A:Accession: S50554
A:Molecule type: DNA
A:Residues: 1-492 <DIE>
A:Cross-references: EMBL:U18796; NID:G603265; PID:G603284; GSPDB:GN00005; MIPS:YER051w
C:Genetics:
A:Gene: MIPS:YER051w
A:Cross-references: SGD:S0000853
A:Map position: 5R

Query Match 34.7%; Score 51; DB 2; Length 492;
Best Local Similarity 33.3%; Pred. No. 46;
Matches 9; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 PHKKEENMLSHLYVSSKDKENISKEN 27
Db 67 PNHKEGYLTAYALITQKGRORNKEN 93
RESULT 12
DENSEMX
malate dehydrogenase (oxaloacetate-decarboxylating) (NADP) (EC 1.1.1.40) - mouse
N:Alternate names: 'malic' enzyme; pyruvic-malic carboxylase
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Jun-2002
C:Accession: A26683; I51951
R:Bagchi, S.; Wise, L.S.; Brown, M.L.; Bregman, D.; Sul, H.S.; Rubin, C.S.
J. Biol. Chem. 262, 1558-1565, 1987
A:Title: Structure and expression of murine malic enzyme mRNA. Differentiation-dependent
A:Reference number: A26683; MUID:87109297; PMID:3805042
A:Accession: A26683
A:Molecule type: mRNA
A:Residues: 1-572 <BAG>
A:Cross-references: GB:J02652; NID:G199762; PIDN:AAA39727.1; PID:G199763
R:Bagchi, S.; Wise, L.S.; Brown, M.L.; Sul, H.S.; Bregman, D.B.; Rubin, C.S.
Ann. N. Y. Acad. Sci. 478, 77-92, 1986
A:Title: Regulation and structure of murine malic enzyme mRNA.
A:Reference number: I51951; MUID:87098422; PMID:3541755
A:Accession: I51951
A:Molecule type: mRNA
A:Residues: 1-572 <BA2>
A:Cross-references: GB:M26756; NID:G199020; PIDN:AAA39489.1; PID:G199021
C:Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)
C:Keywords: homotetramer; NADP; oxidoreductase
F:296-325/Region: beta-alpha-beta NADP nucleotide-binding fold

Query Match 34.7%; Score 51; DB 1; Length 572;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 10 LSHLYVSSKDKENISKEN 27
Db 309 IAHVVVWAMEKEGLSKEN 326
RESULT 13
cnjB protein - Tetrahymena thermophila
C:Species: Tetrahymena thermophila
C>Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 07-Dec-1999
C:Accession: S42136; S42135; S03650
R:Taylor, F.M.; Martindale, D.W.
A:Description: The sequence of S. cerevisiae cosmid 9379, 9581, and lambda clone 4678.
A:Reference number: S42136
A:Accession: S42136

Query Match 34.7%; Score 51; DB 1; Length 572;
Best Local Similarity 50.0%; Pred. No. 54;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 10 LSHLYVSSKDKENISKEN 27
Db 309 IAHVVVWAMEKEGLSKEN 326
RESULT 13
cnjB protein - Tetrahymena thermophila
C:Species: Tetrahymena thermophila
C>Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 07-Dec-1999
C:Accession: S42136; S42135; S03650
R:Taylor, F.M.; Martindale, D.W.
A:Description: The sequence of S. cerevisiae cosmid 9379, 9581, and lambda clone 4678.
A:Reference number: S42136
A:Accession: S42136

A:Molecule type: DNA
A:Residues: 1-1748 <TAY>
A:Cross-references: EMBL:L03710; NID:G161751; PID:G161752
R:Taylor, F.M.; Martindale, D.W.
Nucleic Acids Res. 21, 4610-4614, 1993
A:Title: Retroviral-type zinc fingers and glycine-rich repeats in a protein encoded by cr
A:Reference number: S42135; MUID:94051569; PMID:8233798
A:Accession: S42135
A:Molecule type: DNA
A:Residues: 1164-1174; 1179-1198; 1233-1252; 1285-1293; 1297-1309; 1316-1326; 1331-1341; 1343-13
A:Cross-references: EMBL:L03710
R:Martindale, D.W.; Taylor, F.M.
Nucleic Acids Res. 16, 2189-2201, 1988
A:Title: Multiple introns in a conjugation-specific gene from Tetrahymena thermophila.
A:Reference number: S03650; MUID:88189811; PMID:3357771
A:Accession: S03650
A:Molecule type: DNA
A:Residues: 236-250, 'I', 252-255, 'N', 257-773 <NAR>
A:Cross-references: EMBL:X06462
C:Genetics:
A:Gene: cnjB
A:Genetic code: SGC5
A:Introns: 85/3; 136/1; 157/3; 201/2; 290/2; 327/3; 499/1; 573/2; 607/3; 708/3; 777/3; 81

Query Match 34.4%; Score 50.5; DB 2; Length 1748;
Best Local Similarity 40.0%; Pred. No. 28+02;
Matches 10; Conservative 8; Mismatches 2; Indels 5; Gaps 1;
QY 6 EEMLSHLY-----VSSKDKENISK 25
Db 419 EENLSHIYQNTQLSQERLNLS 443
RESULT 14
T18427
hypothetical protein C0335c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18427
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18427
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3724 <LAW>
A:Cross-references: EMBL:Z98547; NID:el325376; PID:el325379; PIDN:CAB11104.1
C:Genetics:
A:Introns: 307/1; 1545/2
A>Note: C0335c

Query Match 34.4%; Score 50.5; DB 2; Length 1748;
Best Local Similarity 40.0%; Pred. No. 28+02;
Matches 10; Conservative 8; Mismatches 2; Indels 5; Gaps 1;

QY 6 EEMLSHLY-----VSSKDKENISK 25
Db 419 EENLSHIYQNTQLSQERLNLS 443
RESULT 14
T18427
hypothetical protein C0335c - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C:Accession: T18427
R:Lawson, D.; Bowman, S.; Barrell, B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z18935
A:Accession: T18427
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-3724 <LAW>
A:Cross-references: EMBL:Z98547; NID:el325376; PID:el325379; PIDN:CAB11104.1
C:Genetics:
A:Introns: 307/1; 1545/2
A>Note: C0335c

Query Match 34.4%; Score 50.5; DB 2; Length 1748;
Best Local Similarity 40.0%; Pred. No. 28+02;
Matches 10; Conservative 8; Mismatches 2; Indels 5; Gaps 1;

QY 2 EHKKEENML-SHLYVSSKDKENISKEND 28
Db 1791 EHEKDSMLIKHLSVLLKDEKENI-KNNE 1818
RESULT 15
S45428
PET112 protein - yeast (Saccharomyces cerevisiae)

Query Match 34.4%; Score 50.5; DB 2; Length 3724;
Best Local Similarity 55.2%; Pred. No. 4.3e+02;
Matches 16; Conservative 4; Mismatches 6; Indels 3; Gaps 3;

QY 2 EHKKEENML-SHLYVSSKDKENISKEND 28
Db 1791 EHEKDSMLIKHLSVLLKDEKENI-KNNE 1818
RESULT 15
S45428
PET112 protein - yeast (Saccharomyces cerevisiae)

N;Alternate names: protein YBL0724; protein YBL080c
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 21-Jul-2000
C;Accession: S45428; S45821; S41997; S59225; S41792
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
submitted to the EMBL Data Library, May 1994
A;Description: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevisiae chromosome V
A;Reference number: S45387
A;Accession: S45428
A;Molecule type: DNA
A;Residues: 1-541 <OBE>
A;Cross-references: EMBL:X79489; NID:g496661; PID:g496701
R;Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45816
A;Accession: S45821
A;Molecule type: DNA
A;Residues: 1-541 <DOM>
A;Cross-references: EMBL:Z35841; NID:g536128; PID:g536129; GSPDB:GN00002; MIPS:YBL080c
R;Mulero, J.J.; Rosenthal, J.K.; Fox, T.D.
Genet. 25, 299-304, 1994
A;Title: PET112, a Saccharomyces cerevisiae nuclear gene required to maintain rho(+) mitochondria
A;Reference number: S41997; MUID:94363744; PMID:8082172
A;Accession: S41997
A;Molecule type: DNA
A;Residues: 1-414, 'P', 416-541 <MUL>
A;Cross-references: EMBL:L22072; NID:g347492; PIDN:AAC37508.1; PID:g347493
R;Obermaier, B.; Gassenhuber, J.; Piravandi, E.; Domdey, H.
Yeast 11, 1103-1112, 1995
A;Title: Sequence analysis of a 78.6 kb segment of the left end of Saccharomyces cerevisiae chromosome V
A;Reference number: S59184; MUID:96076635; PMID:7502586
A;Accession: S59225
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-541 <OBW>
A;Cross-references: EMBL:X79489; NID:g496661; PIDN:CAA56028.1; PID:g496701
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1994
C;Genetics:
A;Gene: SGD:PET112; MIPS:YBL080c
A;Cross-references: SGD:S0000176; MIPS:YBL080c
A;Map position: 2L
C;Function:
A;Description: involved in mitochondrial gene expression, probably in translation
C;Superfamily: PET112 protein
C;Keywords: mitochondrion; transmembrane protein
P;77-93/Domain: transmembrane #status predicted <TMM>
Query Match 34.0% Score 50; DB 1; Length 541;
Fast Local Similarity 50.0%; Pred. No. 69;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 4 KKEENMLSHLYVSSKQKE 21
DB 204 KKYQLVRLHLHSSGDLE 221
Search completed: September 22, 2004, 18:10:01
Job time : 12.6446 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 5.78512 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: us-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKEENMLSHLYSSKDKENISKEND 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	37.4	1056	1 Z451_MOUSE	Q8C0P7 mus musculus
2	52	35.4	749	1 MADI_YEAST	P40957 saccharomyc
3	52	35.4	993	1 SCPI_MOUSE	Q62209 mus musculus
4	51	34.7	492	1 YE01_YEAST	P40034 saccharomyc
5	51	34.7	572	1 MAOX_MOUSE	P06801 mus musculus
6	50	34.0	541	1 GATB_YEAST	P33893 saccharomyc
7	49	33.3	776	1 ANR5_HUMAN	Q9nu02 homo sapien
8	49	33.3	847	1 GYRA_BUCBP	Q89a83 buchnera ap
9	49	33.3	997	1 SCPI_RAT	Q03410 rattus norv
10	49	33.3	1956	1 ATX1_PLAFA	Q04956 mycoplasma
11	48	32.7	318	1 FMT_MYCPE	Q8ex00 plasmodium
12	48	32.7	504	1 ATPA_CYACA	Q9tm26 cyanidium c
13	48	32.7	845	1 SCPI_MESAU	Q60563 mesocricetu
14	48	32.7	1061	1 Z451_HUMAN	Q9y4e5 homo sapien
15	48	32.7	1679	1 GCC2_MOUSE	Q8c8g3 mus musculus
16	48	32.7	3584	1 NBEA_DROME	Q9w462 drosophila
17	47.5	32.3	646	1 SSRP_ARATH	Q05153 arabidopsis
18	47.5	32.3	889	1 TRK2_YEAST	P28584 saccharomyc
19	47.5	32.3	913	1 UB26_HUMAN	Q9bxu7 homo sapien
20	47.5	32.3	1085	1 CUI7_SCHPO	P24339 schizosacch
21	47	32.0	204	1 YF11_YEAST	P43586 saccharomyc
22	47	32.0	482	1 Y138_METJA	Q57602 methanococc
23	47	32.0	530	1 SYK_METJA	Q57959 methanococc
24	47	32.0	902	1 RNE_BUCAI	P57429 buchnera ap
25	47	32.0	1251	1 YMI1_YEAST	Q04545 saccharomyc
26	46.5	31.6	655	1 YKDA_MYCCA	P45615 mycoplasma
27	46	31.3	190	1 YG31_HAEIN	P44279 haemophilus
28	46	31.3	317	1 TNMD_RAT	Q9esc2 rattus norv
29	46	31.3	354	1 YOIB_CABEL	Q09282 caenorhabdi
30	46	31.3	392	1 YKH7_YEAST	P16081 saccharomyc
31	46	31.3	744	1 NUSC_DAMDI	Q32126 dampiera di
32	46	31.3	794	1 Z148_MOUSE	Q61624 mus musculus
33	46	31.3	794	1 Z148_RAT	Q62806 rattus norv

ALIGNMENTS

RESULT 1
Z451_MOUSE STANDARD; PRT; 1056 AA.
AC Q8C0P7;
DT 15-MAR-2004 (Rel. 43; Created)
DT 15-MAR-2004 (Rel. 43; Last sequence update)
DT 15-MAR-2004 (Rel. 43; Last annotation update)
DE Zinc finger protein 451.
GN ZNF451.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]

20

34 46 31.3 885 1 Y143_HUMAN Q14156 homo sapien
35 46 31.3 1147 1 TEA1_SCHPO P87061 schizosacch
36 46 31.3 4835 1 MDN1_GIALA Q85tcl giardia lam
37 45.5 31.0 293 1 RRP4_HUMAN Q13868 homo sapien
38 45.5 31.0 411 1 HFLK_BUCAP Q8k914 buchnera ap
39 45.5 31.0 573 1 YHDI_SCHPO Q9p619 schizosacch
40 45 30.6 170 1 Y019_BORBU Q1051 borrelia bu
41 45 30.6 216 1 EVGI_MOUSE Q9d981 mus musculus
42 45 30.6 357 1 CADH_POPE P31657 populus del
43 45 30.6 361 1 COL9_ARATH Q02800 arabidopsis
44 45 30.6 422 1 TRD2_METJA Q58759 methanococc
45 45 30.6 424 1 SKP2_MOUSE Q92023 mus musculus

RA SULTANA R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,
Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hiroane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
Yasuishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
CC -1- FUNCTION: May be involved in transcriptional regulation.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
FINGER PROTEINS.
CC -1- SIMILARITY: Contains 12 C2H2-type zinc fingers.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AK030088; BAC26778.1; -.
DR MGD; MGI:2138298; A1596398.
DR InterPro; IPR007087; Znf C2H2.
DR Pfam; PF00096; zf-C2H2; 7.
DR SMART; SM00355; Znf C2H2; 12.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 8.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
DR Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT ZN_FING 169 195 C2H2-TYPE 1.
FT ZN_FING 212 234 C2H2-TYPE 2 (DEGENERATE).
FT ZN_FING 253 277 C2H2-TYPE 3.
FT ZN_FING 315 338 C2H2-TYPE 4 (ATYPICAL).
FT ZN_FING 362 385 C2H2-TYPE 5.
FT ZN_FING 494 517 C2H2-TYPE 6.
FT ZN_FING 527 550 C2H2-TYPE 7.
FT ZN_FING 604 629 C2H2-TYPE 8 (ATYPICAL).
FT ZN_FING 634 657 C2H2-TYPE 9.
FT ZN_FING 665 688 C2H2-TYPE 10.
FT ZN_FING 751 774 C2H2-TYPE 11.
FT ZN_FING 787 810 C2H2-TYPE 12.
SQ SEQUENCE 1056 AA; 120069 MW; 7D8FBC0B50ECA622 CRC64;

Query Match 37.4%; Score 55; DB 1; Length 1056;
Best Local Similarity 38.5%; Pred. No. 15;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 3 HKKEENMLSHLYVSSKDKENISKEND 28

DB 683 HYKHHSDYVVFSEKTKTSIKTEGD 708

RESULT 2

ID MAD1 YEAST STANDARD; PRT; 749 AA.
AC P40357;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Spindle assembly checkpoint component MAD1 (Mitotic MAD1 protein).
GN MAD1 OR YGL086W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1] -
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=96042315; PubMed=7593191;
RA Hardwick K.G., Murray A.W.;
RT "Mad1p, a phosphoprotein component of the spindle assembly checkpoint
in budding yeast";
RL J. Cell Biol. 131:709-720 (1995).
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=S288c;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
chromosome VII";
RL Yeast 13:1077-1090 (1997).
RN [3]

RP INTERACTIONS.
RX MEDLINE=98128031; PubMed=9461437;
RA Hwang L.H., Lau L.F., Smith D.L., Mistrot C.A., Hardwick K.G.,
RA Hwang E.S., Amon A., Murray A.W.;
RT "Budding yeast Cdc20: a target of the spindle checkpoint";
RL Science 279:1041-1044 (1998).
CC -!- FUNCTION: Central component of the spindle assembly checkpoint.

CC -!- SUBUNIT: The spindle checkpoint complex is composed of MAD1, MAD2
CC and MAD3. It interacts with CDC20.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: Becomes hyperphosphorylated when wild-type cells are arrested
CC in mitosis.
CC -!- SIMILARITY: SOME, TO S.POMBE SPBC3D5.04C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U14632; AAA91620.1; -.
DR EMBL; 272608; CAA96791.1; -.
DR PIR; A57276; A57276.
DR GerMOnline; 141134; -.
DR SGD; S0003054; MAD1.
DR GO; GO:0005643; C:nuclear pore; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0007094; P:mitotic spindle checkpoint; IGI.
DR GO; GO:0006913; P:nucleocytoplasmic transport; IGI.
DR InterPro; IPR008672; MAD.
DR Pfam; PF05557; MAD; 1.
KW Mitosis; Coiled coil; Nuclear protein; Phosphorylation.
FT DOMAIN 57 221 COILED COIL (POTENTIAL).
FT DOMAIN 253 324 COILED COIL (POTENTIAL).
FT DOMAIN 330 656 COILED COIL (POTENTIAL).
FT DOMAIN 323 329 POLY-ASP.
FT DOMAIN 330 347 POLY-ASN.
FT DOMAIN 355 363 POLY-ASN.
FT DOMAIN 601 604 POLY-LEU.
SQ SEQUENCE 749 AA; 87651 MW; 312DDE1BE241610D CRC64;

Query Match 35.4%; Score 52; DB 1; Length 749;
Best Local Similarity 42.3%; Pred. No. 26;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 2 EHKKEENMLSHLYVSSKDKENISKEN 27

DB 283 ENEKLNKLSQLHVLESQVENIQLEN 308

RESULT 3

SCPI MOUSE
ID SCPI MOUSE STANDARD; PRT; 993 AA.
AC Q62209; O09205; P70192; Q62329;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Synaptonemal complex protein 1 (SCP-1 protein).
GN SYCP1 OR SCPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBA; TISSUE=Testis;
RX MEDLINE=96004899; PubMed=7548215;
RA Sage J., Martin L., Cuzin F., Rassoulzadegan M.;
RT "cDNA sequence of the murine synaptonemal complex protein 1 (SCPI).";
RL Biochim. Biophys. Acta 1263:258-260 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss; TISSUE=Testis;
RA Kerr S.M., Taggart M.H., Lee M., Cooke H.J.;
RL Submitted (APR-1995) to the EMBL/Genbank/DBSJ databases.
RN [3]
RP SEQUENCE OF 1-149 FROM N.A.
RC STRAIN=C57BL/6;

RA Sage J., Li Y., Martin L., Mattei M.-G., Guenet J.-L., Liu J.G.,
 RL Hoog C., Cuzin F., Rassoulzadegan M.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 95-787 FROM N.A.
 RC STRAIN=ICR; TISSUE=Testis;
 RX Tachida J., Nishina Y., Nozaki M., Uchida K., Nishimune Y.;
 RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Major component of the transverse filaments of
 CC synaptonemal complexes (SCS), formed between homologous
 CC chromosomes during meiotic prophase.
 CC -1- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of
 CC synaptonemal complexes, between lateral elements in the nucleus.
 CC found only where the chromosome cores are synapsed. Its N-terminus
 CC is found towards the centre of the synaptonemal complex while the
 CC C-terminus extends well into the lateral domain of the
 CC synaptonemal complex (By similarity).
 CC -1- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues,
 CC flanked by N- and C-terminal globular domains. The C-terminal
 CC domain has DNA-binding capacity (By similarity).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z38118; CAA86262.1; --
 DR EMBL; L41069; AAG64514.1; ALT_INIT.
 DR EMBL; U62864; AAC53335.1; --
 DR EMBL; U62860; AAC53335.1; JOINED.
 DR EMBL; U62861; AAC53335.1; JOINED.
 DR EMBL; U62862; AAC53335.1; JOINED.
 DR EMBL; U62863; AAC53335.1; JOINED.
 DR EMBL; D88539; BAA13639.1; --
 DR PIR; S49461; S49461.
 DR MGD; MGI:105931; Sycpl.
 DR GO; GO:000795; C:synaptonemal complex; IDA.
 DR InterPro; IPR008827; SCP-1.
 DR Pfam; PF05483; SCP-1; 1.
 KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
 KW DNA-binding; Coiled coil.
 FT DOMAIN 12 97 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 104 815 COILED COIL (POTENTIAL).
 FT DOMAIN 114 117 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 697 700 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 898 901 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 978 986 ARG/LYS-RICH (BASIC).
 FT CONFLICT 527 527 F -> L (IN REF. 2).
 SQ SEQUENCE 993 AA; 115962 MW; 1A4FA790D64FAFE6 CRC64;
 Query Match 35.4%; Score 52; DB 1; Length 993;
 Best Local Similarity 41.4%; Pred. No. 35;
 Matches 12; Conservative 6; Mismatches 7; Indels 4; Gaps 1;
 QY 2 EHKKE-----ENMLSHLYVSSKDKENISKE 26
 DB 253 EYQEVNKNQVSELLQIAEKNRMKD 281
 RESULT 4
 ID YEOL_YEAST STANDARD; PRT; 492 AA.
 AC P40034;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Hypothetical 56.5 kDa protein in CAJ1-HOM3 intergenic region.
 GN YER051W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313264; PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Brenan A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Mossdale D., Nakahara K., Namath A., Norgren R., Oefner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
 RL Nature 387:78-81 (1997).
 CC -1- SIMILARITY: TO C.ELEGANS HYPOTHETICAL PROTEIN T26A5.5.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U18796; AAB64586.1; --
 DR PIR; S50554; S50554.
 DR GERMOnline; I39131; --
 DR SGD; S0000853; YER051W.
 DR InterPro; IPR007113; Cupin sup.
 DR InterPro; IPR003347; TFM_JmJC.
 DR InterPro; IPR001965; Znf_PHD.
 DR Pfam; PF02373; JmJC; 1.
 DR Pfam; PF00628; PHD; 1.
 DR SMART; SM00558; JmJC; 1.
 DR SMART; SM00249; PHD; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 492 AA; 56530 MW; 5CBDD09352131692B CRC64;
 Query Match 34.7%; Score 51; DB 1; Length 492;
 Best Local Similarity 33.3%; Pred. No. 22;
 Matches 9; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
 QY 1 PEHKKEENMLSHLYVSSKDKENISKEN 27
 DB 67 PNKEGEYLTAVALITQKGRQRNKEN 93
 RESULT 5
 ID MAOX_MOUSE STANDARD; PRT; 572 AA.
 AC P06801;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE NADP-dependent malic enzyme (EC 1.1.1.40) (NADP-ME) (Malic enzyme 1).
 GN ME1 OR MOD1 OR MOD-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87109297; PubMed=3805042;
 RA Bagchi S., Wise L.S., Brown M.L., Bregman D., Sul H.S., Rubin C.S.;
 RT "Structure and expression of murine malic enzyme mRNA.
 RT Differentiation-dependent accumulation of two forms of malic enzyme
 RT mRNA in 3T3-L1 cells.";
 RL J. Biol. Chem. 262:1558-1565 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=87098422; PubMed=3541755;

```

RA Bagchi S., Wise L.S., Brown M.L., Sul H.S., Bregman D.B.,
RA Rubin C.S.;
RA "Regulation and structure of murine malic enzyme mRNA.";
RL Ann. N.Y. Acad. Sci. 478:77-92(1986).
CC -!- CATALYTIC ACTIVITY: (S)-malate + NADP(+) = pyruvate + CO(2) +
CC NADPH.
CC
CC -!- SUBUNIT: Homotetramer.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the malic enzymes family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J02652; AAA39727.1; -
CC EMBL; M26756; AAA39489.1; -
CC FIR; A26683; DEMSNX.
CC SWISS-2DPAGE; P06801; MOUSE.
CC MGI; 97043; Mod1.
CC InterPro; IPR001891; Malic_oxred.
CC Pfam; PF00390; malic; 1.
CC Pfam; PF03949; malic_N; 1.
CC PRINTS; PR00072; MALOXRDASE.
CC PROSITE; PS00331; MALIC_ENZYMES; 1.
KW Oxidoreductase; NADP.
FT NP BIND 301 NADP (BY SIMILARITY).
SQ SEQUENCE 572 AA; 63998 MW; 86B1761D04C2B51E CRC64;

Query Match 34.7%; Score 51; DB 1; Length 572;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 10 LSHLYVSSKDKENISKEN 27
DB 309 IAHLVVMAMEKGLSKEN 326

RESULT 6
GATB_YEAST STANDARD; PRT; 541 AA.
AC P33893;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glutamyl-tRNA(Gln) amidotransferase subunit B, mitochondrial
DE precursor (EC 6.3.5.-) (GLU-ADT subunit B) (Cytochrome oxidase
DE assembly factor PET112).
DE PET112 OR YBL080C OR YBL0724.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94363744; PubMed=8082172;
RA Mulero J.J., Rosenthal J.K., Fox T.D.;
RT "PET112, a Saccharomyces cerevisiae nuclear gene required to maintain
RT rho+ mitochondrial DNA.";
RL Curr. Genet. 25:299-304(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=96076635; PubMed=7502586;
RA Obermaier B., Gassenhuber J., Piravandi E., Domdey H.;
RT "Sequence analysis of a 78.6 kb segment of the left end of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 11:1103-1112(1995).
CC -!- FUNCTION: Furnishes a means for formation of correctly charged
CC Gln-tRNA(Gln) through the transamidation of misacylated Glu-

```

```

CC tRNA(Gln) in the mitochondria. The reaction takes place in the
CC presence of glutamine and ATP through an activated gamma-phospho-
CC Glu-tRNA(Gln) (Potential).
CC -!- SUBCELLULAR LOCATION: Mitochondrial (Probable).
CC -!- SIMILARITY: Belongs to the gatB/gate family. GatB subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X79489; CAA56028.1; -
CC EMBL; Z35841; CAA84901.1; -
CC EMBL; L22072; AAC37508.1; -
CC FIR; S45428; S45428.
CC Germonline; 138515; -
CC SGD; S0000176; PET112.
CC GO; GO:0005739; C:mitochondrion; IGI.
CC GO; GO:0009060; P:aerobic respiration; IMP.
CC GO; GO:0007005; P:mitochondrion organization and biogenesis; IMP.
CC GO; GO:0006412; P:protein biosynthesis; IMP.
CC InterPro; IPR004413; GatB.
CC InterPro; IPR006107; GatB.cent.
CC InterPro; IPR006075; GatB_N.
CC InterPro; IPR003789; GatB_Yqey.
CC Pfam; PF01162; GatB; 1.
CC Pfam; PF02934; GatB_N; 1.
CC Pfam; PF02637; GatB_Yqey; 2.
CC TIGRfams; TIGR00133; gatB; 1.
CC PROSITE; PS01234; GATB; 1.
KW Protein biosynthesis; Ligase; Mitochondrion; Transit peptide.
FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
FT CHAIN ? 541 PROBABLE GLUTAMYL-TRNA(GLN)
FT AMIDOTRANSFERASE SUBUNIT B.
FT CONFLICT 415 415 A -> P (IN REF. 1).
SQ SEQUENCE 541 AA; 61842 MW; EE96E82F0F82BBDC CRC64;

Query Match 34.0%; Score 50; DB 1; Length 541;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 KKEENMLSHLYVSSKDKKE 21
DB 204 KKYQNLVRHLHISGDL 221

RESULT 7
ANR5_HUMAN STANDARD; PRT; 776 AA.
ID ANR5_HUMAN
AC Q9NU02; Q9H6Y9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ankyrin repeat domain protein 5.
GN ANKRD5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark L.C., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,

```

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehesvaioho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.B., Martin S.L., McConachie L.J., McWay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Succi C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.D., Symamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.,
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871 (2001).
 RN [2]
 RP SEQUENCE OF 329-776 FROM N.A.
 RS TISSUE=Colon;
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RL "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 8 ANK repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL109754; CAB87618.2;
 CC EMBL; AK025322; BAB15111.1; ALT_INIT.
 CC HSP; P80144; 2MYO.
 CC Genew; HGNC:15803; ANKRD5.
 CC InterPro; IPR002110; ANK.
 CC Pfam; PF00023; ank; 10.
 CC PRINTS; PR01415; ANKYRIN.
 CC SMART; SM00248; ANK; 11.
 CC PROSITE; PS00088; ANK_REPEAT; 6.
 CC PROSITE; PSS0297; ANK_REPEAT_REGION; 1.
 CC ANK repeat; Repeat.
 CC REPEAT 47 76 ANK 1.
 CC REPEAT 184 213 ANK 2.
 CC REPEAT 217 246 ANK 3.
 CC REPEAT 250 279 ANK 4.
 CC REPEAT 524 553 ANK 5.
 CC REPEAT 557 586 ANK 6.
 CC REPEAT 590 619 ANK 7.
 CC REPEAT 623 652 ANK 8.
 CC SEQUENCE 776 AA; 86664 MW; 2F71P35AC4D337B6 CRC64;
 Query Match 33.3%; Score 49; DB 1; Length 776;
 Best Local Similarity 37.9%; Pred. No. 68;
 Matches 11; Conservative 7; Mismatches 9; Indels 2; Gaps 1;
 QY 1 PEKKENMLSHLY--VSSKDKENISKEN 27
 DB 672 PEIKKEELSSIVGPTTSEGGKVKQGN 700
 RESULT 8
 GYRA_BUCBP STANDARD; PRT; 847 AA.
 ID GYRA_BUCBP
 AC Q89AS3;
 DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA gyrase subunit A (EC 5.99.1.3).
 GN GYRA OR BBP169.
 OS Buchnera aphidicola (subsp. Baizongia pistaciae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Buchnera.
 OX NCBI_TaxID=135842;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22436901; PubMed=12522265;
 RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
 RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
 RA Tamames J., Viguera E., Latorre A., Valencia A., Moran F., Moya A.;
 RT "Reductive genome evolution in Buchnera aphidicola.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586 (2003).
 CC -1- FUNCTION: DNA gyrase negatively supercoils closed circular double-
 CC stranded DNA in an ATP-dependent manner and also catalyzes the
 CC interconversion of other topological isomers of double-stranded
 CC DNA rings, including catenanes and knotted rings.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
 CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
 CC enzyme forms an A2B2 tetramer.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE014016; AAO26902.1;
 CC InterPro; IPR006691; DNA_gyraseA_C.
 CC InterPro; IPR002205; DNA_topoisomIV.
 CC Pfam; PF03989; DNA_gyraseA_C; 6.
 CC Pfam; PF00521; DNA_topoisomIV; 1.
 CC Topoisomerase; Isomerase; DNA-binding; Complete proteome.
 CC ACT_SITE 122 122 122 122 DNA_CLEAVAGE (BY SIMILARITY).
 CC SEQUENCE 847 AA; 95829 MW; BCF78EB59DD9939 CRC64;
 Query Match 33.3%; Score 49; DB 1; Length 847;
 Best Local Similarity 40.7%; Pred. No. 74;
 Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;
 QY 2 EHKKEENMLSHLYVSSKDKENISKEND 28
 DB 470 EHKKLISEYKQLFKTSNNLEILKNNN 496
 RESULT 9
 SCPL_RAT STANDARD; PRT; 997 AA.
 ID SCPL_RAT
 AC Q03410;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-JUN-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Synaptonemal complex protein 1 (SCP-1 protein).
 GN SYCP1 OR SCPL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=93099884; PubMed=1464329;
 RA Meuwissen R.L.J., Offenbergh H.H., Dietrich A.J., Riesewijk A.,
 RA Iersel M., Heyting C.;
 RT "A coiled-coil related protein specific for synapsed regions of
 RT meiotic prophase chromosomes.";

EMBO J. 11:5091-5100(1992).

CC -!- FUNCTION: Major component of the transverse filaments of synaptonemal complexes (SCS), formed between homologous chromosomes during meiotic prophase.

CC -!- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of synaptonemal complexes, between lateral elements in the nucleus. Found only where the chromosome cores are synapsed. Its N-terminus is found towards the centre of the synaptonemal complex while the C-terminus extends well into the lateral domain of the synaptonemal complex.

CC -!- TISSUE SPECIFICITY: Testis.

CC -!- DEVELOPMENTAL STAGE: Expressed exclusively in meiotic prophase cells.

CC -!- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues, flanked by N- and C-terminal globular domains. The C-terminal domain has DNA-binding capacity.

CC -!- CAUTION: This is a conceptual translation; a frameshift was corrected in position 6 to maximize the similarity with the other species SYCP1 sequences.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X67805; CAA48006.1; ALT_FRAME.

DR InterPro; IPR008827; SCP-1.

DR Pfam; PF05483; SCP-1; 1.

KW Nuclear protein; Meiosis; Cell division; Phosphorylation; DNA-binding; Coiled coil.

FT DOMAIN 108 819 COILED COIL (POTENTIAL).

FT DOMAIN 118 121 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 701 704 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 902 905 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).

FT DOMAIN 982 990 ARG/LYS-RICH (BASIC).

FT SEQUENCE 997 AA; 116511 MW; 229D59823FD684BE CRC64;

SQ

Query Match 33.3%; Score 49; DB 1; Length 997;
Best Local Similarity 41.4%; Pred. No. 89;
Matches 12; Conservative 6; Mismatches 7; Indels 4; Gaps 1;

QY 2 EHKKE-----ENMLSHYVSSKDKENISKE 26
D 257 EYQKEVNNKENQVSLLIQSTENKMKD 285

RESULT 10

ATX1_PLAFA

ID ATX1_PLAFA STANDARD; PRT; 1956 AA.

AC Q04956;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Probable cation-transporting ATPase 1 (EC 3.6.3.-).

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=19/96;

RX MEDLINE=93132070; PubMed=8421054;

RA Krihnha S., Cowan G., Meade J.C., Wells R.A., Stringer J.R., Robson K.J.;

RT "A family of cation ATPase-like molecules from Plasmodium falciparum";

RL J. Cell Biol. 120:385-398(1993).

CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). Subfamily V.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X65738; CAA46646.1; -.

DR InterPro; IPR001757; ATPase_E1-E2.

DR Pfam; PF00122; E1-E2_ATPase; 1.

DR TIGSFAM6; TIGR01494; ATPase_P-type; 7.

DR PROSITE; PS00154; ATPASE_E1-E2; 1.

KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.

FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 36 58 POTENTIAL.

FT DOMAIN 59 61 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 62 80 POTENTIAL.

FT DOMAIN 81 407 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 408 427 POTENTIAL.

FT DOMAIN 428 440 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 441 462 POTENTIAL.

FT DOMAIN 463 1818 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 1819 1837 POTENTIAL.

FT DOMAIN 1838 1845 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1846 1863 POTENTIAL.

FT DOMAIN 1864 1881 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 1882 1905 POTENTIAL.

FT DOMAIN 1906 1928 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1929 1952 POTENTIAL.

FT DOMAIN 1953 1956 CYTOPLASMIC (POTENTIAL).

FT MOD_RES 496 496 PHOSPHORYLATION (BY SIMILARITY).

FT METAL 1760 1760 MAGNESIUM (BY SIMILARITY).

FT METAL 1764 1764 MAGNESIUM (BY SIMILARITY).

FT DOMAIN 246 251 POLY-ASN.

FT DOMAIN 252 256 POLY-LYS.

FT DOMAIN 937 941 POLY-ASN.

FT DOMAIN 1344 1347 POLY-LYS.

FT DOMAIN 1363 1372 POLY-ASN.

FT DOMAIN 1680 1684 POLY-ASN.

SQ SEQUENCE 1956 AA; 230285 MW; AE708AAE99009335 CRC64;

Query Match 33.3%; Score 49; DB 1; Length 1956;
Best Local Similarity 36.1%; Pred. No. 1.9e+02;
Matches 13; Conservative 6; Mismatches 7; Indels 10; Gaps 2;

QY 2 EHKKEENMLSDFIHKEMNTESIHSKDNMIHNKN 549
D 514 QNKKNKMLSDFIHKEMNTESIHSKDNMIHNKN 549

RESULT 11

FMT_MYCPE

ID FMT_MYCPE STANDARD; PRT; 318 AA.

AC Q8EX00;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Methionyl-tRNA formyltransferase (EC 2.1.2.9).

GN FMT OR MYPE500.

OS Mycoplasma penetrans.

OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.

OX NCBI_TaxID=28227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HF-2;

RX MEDLINE=22354719; PubMed=12466555;

RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K., Yashino C., Horino A., Shiba T., Sasaki T., Hattori M.;

RT "The complete genomic sequence of Mycoplasma penetrans, an intracellular bacterial pathogen in humans.";

CC has three main subunits: a, b and c.
CC -1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -1- SIMILARITY: Belongs to the ATPase alpha/beta chains family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF02186; AAF13005.1; -
CC InterPro: IPR005294; ATP_synthF1_alph.
CC InterPro: IPR000793; ATPase_a/b_C.
CC InterPro: IPR000194; ATPase_a/bCentre.
CC InterPro: IPR004100; ATPase_a/bN.
CC InterPro: IPR000790; ATPase_a_C.
CC InterPro: IPR009005; F1_ATPase_a/bN.
CC Pfam: PF00006; ATP-synt_ab; 1.
CC Pfam: PF00306; ATP-synt_ab_C; 1.
CC Pfam: PF02874; ATP-synt_ab_N; 1.
CC ProDom: PD001099; ATPase_aC; 1.
CC TIGRFAMs: TIGR00962; atpA; 1.
CC PROSITE: PS00152; ATPASE_ALPHA_BETA; 1.
CC ATP synthesis; Chloroplast; Thylakoid; Membrane; CF(1);
CC ATP-binding; Hydrolase; Hydrogen ion transport.
CC NP_BIND 170 177 ATP (BY SIMILARITY).
CC ACT_SITE 363 363 BY SIMILARITY.
CC SEQUENCE 504 AA; 54994 MW; EB714398D887B9E0 CRC64;

CC Query Match 32.7%; Score 48; DB 1; Length 504;
CC Best Local Similarity 37.9%; Pred. No. 57;
CC Matches 11; Conservative 6; Mismatches 8; Indels 4; Gaps 1;

CC QY 4 KKEENWLS-----HLVSSKDKENISKEND 28
CC ||| : ||| : ||| : ||| : ||| :
CC DB 459 KKLQELCLNSYPHFYEAIKESQLSKENE 487

CC RESULT 13
CC SCPL_MESAU
CC ID SCPL_MESAU STANDARD; PRT; 845 AA.
CC AC Q60563;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Synaptonemal complex protein 1 (SCP-1 protein) (Meiotic chromosome
CC DE synaptic protein) (Fragment).
CC OS SCPL OR SYNI.
CC OS Mesocricetus auratus (Golden hamster).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
CC OC Mesocricetus.
CC OX NCBI_TaxID=10036;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Testis;
CC RX MEDLINE=95181577; PubMed=7876343;
CC RA Dobson M.J., Pearlman R.E., Karaiskakis A., Spyropoulos B.,
CC RA Moens P.B.;
CC RT "Synaptonemal complex proteins: occurrence, epitope mapping and
CC RT chromosome disjunction";
CC RL J. Cell Sci. 107:2749-2760(1994).
CC -1- FUNCTION: Major component of the transverse filaments of
CC synaptonemal complexes (SCs), formed between homologous
CC chromosomes during meiotic prophase (By similarity). Has non-
CC specific DNA binding capability.
CC -1- SUBCELLULAR LOCATION: Nuclear. In tripartite segments of
CC synaptonemal complexes, between lateral elements in the nucleus.
CC Found only where the chromosome cores are synapsed. Its N-terminus
CC is found towards the centre of the synaptonemal complex while the
CC C-terminus extends well into the lateral domain of the

```

CC synaptonemal complex.
CC -!- DOMAIN: Consists of an alpha-helical stretch of 700 AA residues,
CC flanked by N- and C-terminal globular domains. The C-terminal
CC domain has DNA-binding capacity (By similarity).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32978; AAC2039.1; -.
DR PIR; I48176; I48176.
DR InterPro; IPR008827; SCP-1.
DR Pfam; PF05483; SCP-1.
DR KX Nuclear protein; Meiosis; Cell division; Phosphorylation;
DR DNA-binding; Coiled coil.
DR NON_TER 1
DR DOMAIN <1 672 COILED COIL (POTENTIAL).
DR DOMAIN 553 556 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
DR DOMAIN 753 756 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
DR DOMAIN 830 838 ARG/LYS-RICH (BASIC).
DR SQ SEQUENCE 845 AA; D7F28873C824C6A8 CRC64;
Query Match 32.7%; Score 48; DB 1; Length 845;
Best Local Similarity 41.4%; Pred. No. 1e+02;
Matches 12; Conservative 5; Mismatches 8; Indels 4; Gaps 1;
QY 2 EHKKE----ENMLSHLYSSKDKENISKE 26
DB 109 EYKKEVNDRENQVSLLIQTEKENKMKD 137
:||||| :||| :||| :||| :
:||||| :||| :||| :||| :
RESULT 14
Z451 HUMAN STANDARD; PRT; 1061 AA.
AC Q9Y4E5; Q8N380; Q8TD15; Q9NQM1;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Zinc finger protein 451 (Coactivator for steroid receptors).
GN ZNF451 OR COACTOR OR KIA00576.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Heldens I.M., Decherling K.J.;
RT "Isolation of a novel coactivator for steroid receptors that alters
RT the intrinsic activity of the estrogen receptor alpha liganded with
RT SERMs."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 5:31-39(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22935763; PubMed=14574404;
RA Mungall A.J., Palmer S.A., Sims S.K., Edwards C.A., Ashurst J.L.,
RA Wilming L., Jones M.C., Horton R., Hunt S.E., Scott C.E.,
RA Gilbert J.G.R., Clamp M.B., Bethel G., Milne S., Ainscough R.,
RA Almeida J.P., Ambrose K.D., Andrews T.D., Ashwell R.I.S.,
RA Babbage A.K., Bagguley C.L., Bailey J., Banerjee R., Barker D.J.,

```

```

RA Barlow K.F., Bates K., Beare D.M., Beasley H., Beasley O., Bird C.P.,
RA Blakey S., Bray-Allen S., Brook J., Brown A.J., Brown J.Y.,
RA Burford D.C., Burrill W., Burton J., Carder C., Carter N.P.,
RA Chapman J.C., Clark S.Y., Clark G., Clee C.M., Clegg S., Cobley V.,
RA Collier R.E., Collins J.E., Colman L.K., Corby N.R., Coville G.J.,
RA Culley K.M., Dhami P., Davies J., Dunn M., Earthrowl M.E.,
RA Ellington A.E., Evans K.A., Faulkner L., Francis M.D., Frankish A.,
RA Frankland J., French L., Garner P., Garnett J., Ghori M.J.,
RA Gilby L.M., Gillson C.J., Glithero R.J., Grafham D.V., Grant M.,
RA Gribble S., Griffiths C., Griffiths M.N.D., Hall R., Halls K.S.,
RA Hammond S., Harley J.L., Hart E.A., Heath P.D., Heathcote R.,
RA Holmes S.J., Howden P.J., Howe K.L., Howell G.R., Huckle E.,
RA Humphray S.J., Humphries M.D., Hunt A.R., Johnson C.M., Joy A.A.,
RA Kay M., Keenan S.J., Kimberley A.M., King A., Laird G.K., Langford C.,
RA Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C.R., Lloyd D.M.,
RA Loveland J.E., Lovell J., Martin S., Mashreghi-Mohammadi M.,
RA Maslen G.L., Matthews L., McCann O.T., McLaren S.J., McLeay K.,
RA McMurray A., Moore M.J.F., Mullikin J.C., Niblett D., Nickerson T.,
RA Novik K.L., Oliver K., Overton-Larty E.K., Parker A., Patel R.,
RA Pearce A.V., Peck A.I., Phillimore B.J.C.T., Phillips S., Plumb R.W.,
RA Porter K.M., Ramsey Y., Randy S.A., Rice C.M., Ross M.T., Searle S.M.,
RA Sehra H.K., Sheridan E., Skuce C.D., Smith S., Smith M., Spraggon L.,
RA Squares S.I., Steward C.A., Sycamore N., Tamlyn-Hall G., Tester J.,
RA Theaker A.J., Thomas D.W., Thorpe A., Tracey A., Tromans A., Tubby B.,
RA Wall M., Wallig J.M., West A.P., White S.S., Whitehead S.L.,
RA Whitaker H., Wild A., Willey D.J., Wilmer T.E., Wood J.M., Wray P.W.,
RA Wyatt J.C., Young L., Younger R.M., Bentley D.R., Coulson A.,
RA Durbin R., Hubbard T., Sulston J.E., Dunham I., Rogers J., Beck S.;
RT "The DNA sequence and analysis of human chromosome 6."
RL Nature 425:805-811(2003).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be involved in transcriptional regulation.
CC Coactivator for steroid receptors.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9Y4E5-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9Y4E5-2; Sequence=VSP_008624;
CC Name=3;
CC IsoId=Q9Y4E5-3; Sequence=VSP_008625; VSP_008626;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 11 C2H2-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


DR	ENBL;	BC027339;	AAH27339.1;	ALT_INIT.
DR	ENBL;	AK011206;	BAB27466.2;	-
DR	ENBL;	AK045701;	BAC32463.1;	-
DR	ENBL;	AB093232;	BAC41416.1;	-
DR	PR;	PT0849;	PT0712.	-
DR	MGD;	MGI1917547;	2600014C01Rik.	-
DR	InterPro;	IPR002037;	GRIP domain.	-
DR	InterPro;	IPR002017;	Spectrin.	-
DR	Pfam;	PF01465;	GRIP; 1.	-
DR	PROSITE;	PS50913;	GRIP; 1.	-
DR	GO	golgi stack;	Coiled coil.	-
KW	DOMAIN	31	1613	COILED COIL
FT	DOMAIN	83	501	GLU-RICH
FT	DOMAIN	1604	1654	GRIP.
FT	DOMAIN	1604	1654	GRIP.
FT	CONFLICT	440	440	I -> M (
SQ	SEQUENCE	1679	AA;	194443 MW; 652F

Query Match	32.7%	Score 48;	DB 1;	Length 1679;
Local Similarity	36.4%	Pred. No. 2.1e+02;		
Conservative	8;	Mismatches 7;	Indels 0;	Gaps 0;

QY 3 HKKEENMLSHLYVSSKDKENIS 24
 |::| :|| :|| :|| :|| :|| ::
D6 428 HQKEVSELSETFISGSEKEKLA 449

Search completed: September 22, 2004, 18:07:00
Job time : 6.78512 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:02:42 ; Search time 10.2645 Seconds
(without alignments)
253.026 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKBASSVDYILGWFGGVPFHKKEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	61	40.7	1238	T34929	hypothetical prote
2	57	38.0	382	1 SYECCS	carbamoyl-phosphat
3	57	38.0	382	2 B85484	carbamoyl-phosphat
4	57	38.0	382	2 C90633	carbamoyl-phosphat
5	55	36.7	479	2 C64655	hypothetical prote
6	55	36.7	479	2 H71942	probable outer mem
7	53	35.3	682	2 T47473	receptor-like prot
8	52.5	35.0	181	2 S78336	conserved hypothet
9	52.5	35.0	602	2 G81195	aspartyl-tRNA synt
10	52.5	35.0	602	2 F81931	aspartate-tRNA lig
11	52	34.7	858	2 T18946	probable phospholi
12	51	34.0	271	2 E70240	pfs protein (pfs)
13	50.5	33.7	588	2 E89945	aspartyl-tRNA synt
14	50	33.3	577	2 A64597	aspartate-tRNA lig
15	50	33.3	579	2 C71916	aspartate-tRNA lig
16	49.5	33.0	168	2 T09333	module specific pr
17	49.5	33.0	2025	2 T21588	hypothetical prote
18	49	32.7	587	2 T48582	auxin-regulated pr
19	49	32.7	613	2 T47483	receptor like prot
20	49	32.7	688	2 T48176	receptor like prot
21	48.5	32.3	420	2 AG1385	B. subtilis YvIB p
22	48	32.0	310	2 C83800	ferrochelatase hem
23	48	32.0	557	2 D83478	gamma-glutamyltran
24	48	32.0	564	2 F71975	methyl-accepting c
25	48	32.0	565	2 G64532	methyl-accepting c
26	48	32.0	567	2 E82956	probable potassium
27	48	32.0	682	2 T48175	receptor like prot
28	47	31.3	416	2 I61744	casein kinase I-ep
29	47	31.3	591	2 AG1264	aspartyl-tRNA synt

aspartyl-tRNA synt
maltodextrin glyco
probable amine oxi
amine oxidase-like
fimbrial usher pro
hypothetical prote
isoleucyl-tRNA syn
hypothetical prote
hypothetical prote
xyloglucan endo-1,
acid phosphatase (
receptor like prot
xylan 1,4-beta-xy
geranylgeranyl pyr
geranyl transxans
nitrilotriacetate

ALIGNMENTS

RESULT 1

T34929
hypothetical protein SC3F9.07 SC3F9.07 - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
C/Accession: T34929
R/Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z21562
A/Accession: T34929
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1238 <SEE>
A/Cross-references: EMBL:AL023862; PIDN:CAAL19630.1; GSPDB:GN00070; SCOEDB:SC3F9.07
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC3F9.07

Query Match 40.7%; Score 61; DB 2; Length 1238;
Best Local Similarity 44.8%; Pred. No. 1.5;
Matches 13; Conservative 2; Mismatches 8; Indels 6; Gaps 1;

QY 1 AKBASSVDYILGWFGG-----VPEHK 23
DB 531 AKDADFDLTGWDGDKGKGLTTHK 559

RESULT 2

SYECCS
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) small chain [validated]
N/Alternate names: carbamoyl-phosphate synthetase glutamine chain
C/Species: Escherichia coli
C/Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 01-Mar-2002
C/Accession: A01128; S40555; H64723
R/Lusty, C.J.
submitted to the Protein Sequence Database, April 1984
A/Reference number: A01128
A/Accession: A01128
A/Molecule type: protein
A/Residues: 1-382 <LUS>
R/Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobuchi, T.
submitted to the EMBL Data Library, December 1992
A/Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2
A/Reference number: S40531
A/Accession: S40555
A/Molecule type: DNA
A/Residues: 1, 2-382 <YUR>
A/Cross-references: EMBL:DJ0483
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

R/Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
C/Accession: G81195
R/Hickey, E.K.; Hafe, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000

A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A/Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A/Reference number: AB1000; MUID:20175755; PMID:10710307

A/Accession: G81195
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-602 <TST>
A/Cross-references: GB:AE002403; GB:AE002098; NID:g7225688; PIDN:AAF40903.1; PID:g7225693
A/Experimental source: serogroup B, strain MC58
C/Genetics:
A/Gene: NME0466
C/Superfamily: lysine-tRNA ligase

Query Match 35.0%; Score 52.5; DB 2; Length 602;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 4 ASSYDYIL-GWEFGGG 18
| : | | : | | | | |
Db 480 ARAYDMVLNGWEIGGG 495

RESULT 10
F81831.
aspartate-tRNA ligase (EC 6.1.1.12) NMA2019 [imported] - *Neisseria meningitidis* (strain)
C/Species: *Neisseria meningitidis*
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 03-Jun-2002
C/Accession: F81831
R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morell,
J.; Holroyd, S.; Jørgels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.
A/Reference number: AB1775; MUID:2022556; PMID:10761919
A/Accession: F81831
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-602 <PAR>
A/Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85238.1; PID:g738064f
A/Experimental source: serogroup A, strain Z2491
C/Genetics:
A/Gene: aspS; NMA2019
C/Superfamily: lysine-tRNA ligase
C/Keywords: ligase

Query Match 35.0%; Score 52.5; DB 2; Length 602;
Best Local Similarity 62.5%; Pred. No. 12;
Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 4 ASSYDYIL-GWEFGGG 18
| : | | : | | | | |
Db 480 ARAYDMVLNGWEIGGG 495

RESULT 11
T18946
probable phospholipase activating protein C05C10.6 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C/Accession: T18946; T24252
R/Matthews, P.
Submitted to the EMBL Data Library, February 1995
A/Reference number: Z19049
A/Accession: T18946
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-858 <WIL>
A/Cross-references: EMBL:Z48178; PIDN:CAA88206.1; GSPDB:GN00020; CESP:C05C10.6

A:Experimental source: clone C05C10
R:Wilkinson, J.
Submitted to the EMBL Data Library, October 1995
A:Reference number: Z19863
A:Accession: T24252
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-858 <W12>
A:Cross-references: EMBL:266515; PIDN:CAA91354.1; GSPDB:GN000020; CESP:C05C10.6
A:Experimental source: clone R53
C:Genetics:
A:Gene: CESP:C05C10.6
A:Map position: 2
A:Introns: 15/3; 120/1; 155/3; 407/3; 513/1; 549/1; 593/3; 711/2; 786/3; 821/3

Query Match 34.7%; Score 52; DB 2; Length 858;
Best Local Similarity 40.0%; Pred. No. 21;
Matches 10; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

3 EASSYDIILGWFFGGVPEHKKEEN 27
: ||| ||| ||| ||| : : : :
Db 427 DGKEYDIALGVNFGKEPDKQPFNF 451

RESULT 12
E70240
pfs protein (pfs) homolog - Lyme disease spirochete plasmid I/1p28-4
C:Species: Borrelia burgdorferi (lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
A:Accession: E70240
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: E70240
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-271 <KLE>
A:Cross-references: GB:AE000789; NID:g2690079; PIDN:AAC66190.1; PID:g2690087; TIGR:BB106
A:Experimental source: strain B31
C:Genetics:
A:Genome: plasmid
C:Superfamily: Escherichia coli pfs protein

Query Match 34.0%; Score 51; DB 2; Length 271;
Best Local Similarity 50.0%; Pred. No. 8.7;
Matches 13; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

3 EASSYDIIL---GWFFGGVPEHKKE 25
: ||| ||| ||| ||| ||| :
Db 137 ETTSYDFDLHRRFYEI-GHVPEHPKK 161

RESULT 13
E89945
aspartyl-tRNA synthetase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
A:Accession: E89945
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89945
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-588 <KUR>
A:Cross-references: GB:BA000018; PID:g13701428; PIDN:BA842722.1; GSPDB:GN00149

A:Experimental source: strain N315
C:Genetics:
A:Gene: asps
C:Superfamily: lysine-tRNA ligase

Query Match 33.7%; Score 50.5; DB 2; Length 588;
Best Local Similarity 58.8%; Pred. No. 23;
Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Qy 3 EASSYDIIL-GWFFGG 18
||: ||| :|: |||
Db 473 EANAYDIVLNGVELGGG 489

RESULT 14
A64597
aspartate-tRNA ligase (EC 6.1.1.12) - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 03-Jun-2002
A:Accession: A64597
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.N
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: A64597
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-577 <TOM>
A:Cross-references: GB:AE000576; GB:AE000511; NID:g2313736; PIDN:AAD07682.1; PID:g2313733;
A:Experimental source: strain 26695
C:Genetics:
A:Gene: asps
C:Function:
A:Description: activates amino acid and transfers it to specific tRNA molecule
A:Pathway: protein biosynthesis
C:Superfamily: lysine-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 33.3%; Score 50; DB 2; Length 577;
Best Local Similarity 52.0%; Pred. No. 27;
Matches 13; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

Qy 3 EASSYDIIL-GWFFGG-VPEHKKE 25
||: ||| :|: |||
Db 462 EAHAYDVVLNGVELGGGSIRIHKEE 486

RESULT 15
C71916
aspartate-tRNA ligase (EC 6.1.1.12) - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 03-Jun-2002
A:Accession: C71916
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; J
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: C71916
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-579 <ARN>
A:Cross-references: GB:AE001489; GB:AE001439; NID:g4155102; PIDN:AAD06143.1; PID:g4155111
A:Experimental source: strain J99
C:Genetics:
A:Gene: asps
C:Superfamily: lysine-tRNA ligase
C:Keywords: ligase


```
Query Match      33.3%; Score 50; DB 2; Length 579;
Best Local Similarity 52.0%; Pred. NO. 27;
Matches 13; Conservative 4; Mismatches 6; Indels 2; Gaps 2;
```

Search completed: September 22, 2004, 18:09:59
Job time : 11.2645 secs

The Page Mark (center)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 5.57851 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKEASSYDYLWEGFGVPEHKKEEN 27

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Number of hits satisfying chosen parameters: 141681

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	38.0	382	1 CARA_ECOLI	P00907 escherichia
2	56.5	37.7	605	1 SYD_BALSO	Q9270 raietonia s
3	55	36.7	382	1 CARA_ECOL6	Q81b1 escherichia
4	52.5	35.0	181	1 YCF4_ODOSI	P49526 odontellia s
5	52.5	35.0	602	1 SYD_NEIMA	Q9jt23 neisseria m
6	52.5	35.0	602	1 SYD_NEIMB	Q9k0u5 neisseria m
7	50.5	33.7	598	1 SYD_STAMP	Q99t19 staphylococ
8	50	33.3	577	1 SYD_HELPY	P56459 helicobacte
9	50	33.3	579	1 SYD_HELPJ	Q9z119 helicobacte
10	50	33.3	598	1 SYD_LACPL	Q88vq8 lactobacill
11	48	32.0	310	1 HEMZ_BACHD	Q9kdk9 bacillus ha
12	47.5	31.7	598	1 SYD_STAPP	Q8c99 staphylococ
13	47	31.3	416	1 KCLF_HUMAN	P49674 homo sapien
14	47	31.3	591	1 SYD_LISIN	Q92b74 listeria in
15	47	31.3	591	1 SYD_LISMO	Q8y709 listeria mo
16	47	31.3	835	1 RNFC_PASMU	Q9cnp2 pasteurella
17	46.5	31.0	412	1 PHOA_PENCH	P37274 penicillium
18	46.5	31.0	523	1 C756_CAMME	O04773 campanula m
19	46	30.7	234	1 HISA_STAPP	Q8c93 staphylococ
20	46	30.7	259	1 K10E_STRGR	P29785 streptomyce
21	46	30.7	416	1 KCLF_MOUSE	Q9jmk2 mus musculu
22	46	30.7	439	1 YDDW_ECOLI	P76130 escherichia
23	46	30.7	943	1 SYI_PSEPL	P18330 pseudomonas
24	45.5	30.3	516	1 ACHI_MANSE	P91766 manduca sex
25	45.5	30.3	522	1 SYD_EACSU	Q32038 bacillus su
26	45	30.0	335	1 Y719_ANASP	Q05067 anabaena sp
27	45	30.0	342	1 QUEA_CAMJE	Q9pht9 campylobact
28	45	30.0	359	1 K6P1_ANASP	O8yk93 anabaena sp
29	45	30.0	537	1 SYE_TREPA	O83679 treponema p
30	45	30.0	598	1 SYD_YERPE	Q8zev0 yersinia pe
31	45	30.0	740	1 RNFC_ECO57	P58324 escherichia
32	45	30.0	740	1 RNFC_ECOLI	P77611 escherichia
33	45	30.0	3712	1 ACVS_CEPAC	P25464 cephalospor

34	44.5	29.7	533	1 QUTD_EMENI	P15325 emericella
35	44.5	29.7	595	1 SYD_BACHD	Q9kdg1 bacillus ha
36	44.5	29.7	678	1 YIHQ_ECOLI	P32138 escherichia
37	44.5	29.7	753	1 EPLI_MOUSE	Q9erg0 mus musculu
38	44	29.3	181	1 RLS_METVA	P14029 methanococc
39	44	29.3	301	1 HKDD_CHICK	P24344 gallus gall
40	44	29.3	303	1 PQQB_ACICA	P07779 acinetobact
41	44	29.3	312	1 HEMZ_OCEIH	Q8erx9 oceanobacil
42	44	29.3	382	1 CARA_SALTI	Q8z918 salmonella
43	44	29.3	382	1 CARA_SALTY	P14845 salmonella
44	44	29.3	397	1 O22H_DROME	P81910 drosophila
45	44	29.3	494	1 SYE_FSEAE	Q9xc16 pseudomonas

ALIGNMENTS

RESULT 1

CARA_ECOLI	18	STANDARD;	PRT;	382 AA.
ID	CARA_ECOLI			
AC	P00907;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase glutamine chain).			
GN	CARA OR PYRA OR B0032 OR Z0037 OR ECS0035.			
OS	Escherichia coli, and			
OS	Escherichia coli, O157:H7.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=562, 83334;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84248072; PubMed=6330744;			
RA	Piette J., Nyumoya H., Lusty C.J., Cunin R., Weyens G., Crabbeel M., Charlier D.R.M., Glandsdorff N., Pierard A.;			
RT	"DNA sequence of the carA gene and the control region of carAB: tandem promoters, respectively controlled by arginine and the pyrimidines, regulate the synthesis of carbamoyl-phosphate synthetase in Escherichia coli K-12.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:4134-4138(1984).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84248073; PubMed=6377309;			
RA	Bouvier J., Patte J.-C., Stragier P.;			
RT	"Multiple regulatory signals in the control region of the Escherichia coli carAB operon.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 81:4139-4143(1984).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92334977; PubMed=1630901;			
RA	Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N., Isono K., Mizobuchi K., Nakata A.;			
RT	"Systematic sequencing of the Escherichia coli genome: analysis of the 0-2.4 min region.";			
RL	Nucleic Acids Res. 20:3305-3308(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=K12 / MG1655;			
RA	MEDLINE=97426617; PubMed=9278503;			
RT	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;			
RL	"The complete genome sequence of Escherichia coli K-12.";			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN=O157:H7 / EDL933 / ATCC 700927;			
RA	MEDLINE=21074935; PubMed=11206551;			
RT	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,			

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobstein E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,
RA Apodaca J., Anantharaman I.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RL "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Okteubo E., Sakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RP SEQUENCE OF 361-382 FROM N.A.
RX MEDLINE=83273669; PubMed=6308632;
RA Nyunoya H., Lusty C.J.;
RA "The carb gene of *Escherichia coli*: a duplicated gene coding for the
RT large subunit of carbamoyl-phosphate synthetase.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4629-4633(1983).
RN [8]
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=93096050; PubMed=1334233;
RA Wang M.X., Church G.M.;
RA "A whole genome approach to in vivo DNA-protein interactions in *E.*
RT coli.";
RL Nature 360:606-610(1992).
RN [9]
RP SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of *Escherichia coli* K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=97317071; PubMed=9174345;
RA Thoden J.B., Holden H.M., Wesenberg G., Raushel F.M., Rayment I.;
RA "Structure of carbamoyl phosphate synthetase: a journey of 96 A from
RT substrate to product.";
RL Biochemistry 36:6305-6316(1997).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=98301363; PubMed=9636022;
RA Thoden J.B., Miran S.G., Phillips J.C., Howard A.J., Raushel F.M.,
RA Holden H.M.;
RT "Carbamoyl phosphate synthetase: caught in the act of glutamine
RT hydrolysis.";
RL Biochemistry 37:8825-8831(1998).
RN [12]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99190825; PubMed=10089390;
RA Thoden J.B., Raushel F.M., Benning M.M., Rayment I., Holden H.M.;
RT "The structure of carbamoyl phosphate synthetase determined to 2.1-A
RT resolution.";
RL Acta Crystallogr. D 55:8-24(1999).
RN [13]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99155207; PubMed=10029528;
RA Thoden J.B., Wesenberg G., Raushel F.M., Holden H.M.;
RT "Carbamoyl phosphate synthetase: closure of the B-domain as a result
RT of nucleotide binding.";
RL Biochemistry 38:2347-2357(1999).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=20056035; PubMed=10587438;
RA Thoden J.B., Huang X., Raushel F.M., Holden H.M.;
RT "The small subunit of carbamoyl phosphate synthetase: snapshots along
RT the reaction pathway.";
RL Biochemistry 38:16158-16166(1999).
RN [15]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=99357782; PubMed=10428826;
RA Thoden J.B., Raushel F.M., Wesenberg G., Holden H.M.;
RT "The binding of inosine monophosphate to *Escherichia coli* carbamoyl
RT phosphate synthetase.";
RL J. Biol. Chem. 274:22502-22507(1999).
CC -!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -!- PATHWAY: Arginine biosynthesis.
CC -!- PATHWAY: Pyrimidine biosynthesis; first step.
CC -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate.
CC Tetramer of heterodimers (alpha,beta)4.
CC -!- SIMILARITY: Belongs to the carA family.
CC -!- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; D10483; BAB96601.1; -.
DR EMBL; AE000113; AAC73143.1; -.
DR EMBL; AE005180; AAG54334.1; -.
DR EMBL; AF002550; BAB33458.1; -.
DR EMBL; AF001717; CAA49615.1; -.
DR PIR; A01128; SYECCS.
DR PIR; B85484; B85484.
DR PIR; C90633; C90633.
DR PDB; 1DUB; 17-JUN-98.
DR PDB; 1A9X; 21-OCT-98.
DR PDB; 1BXR; 20-APR-99.
DR PDB; 1CE8; 26-JUL-99.
DR PDB; 1C30; 10-DEC-99.
DR PDB; 1C30; 10-DEC-99.
DR PDB; 1CS0; 10-DEC-99.
DR PDB; 1KEE; 21-DEC-01.
DR PDB; 1M6V; 13-NOV-02.
DR SWISS-2DPAGE; P00907; COL1.
DR ECO2DBASE; G041.4; 6TH EDITION.
DR EcoGene; EG10134; carA.
DR HAMAP; MF_01209; -; 1.
DR InterPro; IPR006274; CarA_synth_small.
DR InterPro; IPR001317; CP_synthGATase.
DR InterPro; IPR002474; CP_synthsmall.
DR InterPro; IPR000991; GATase 1.
DR Pfam; PF00988; CPSase_sm_chain; 1.
DR Pfam; PF00117; GATase; 1.
DR PRINTS; P00099; CPSGATASE.
DR PRINTS; P00096; GATASE.
DR TIGRFAMs; TIGR01368; CPSaseIismall; 1.
DR PROSITE; PS00442; GATASE TYPE I; 1.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase;
KW Glutamine amidotransferase; 3D-structure; Complete proteome.
FT DOMAIN 1 189 CFSASE.
FT DOMAIN 190 382 GLUTAMINE AMIDOTRANSFERASE.
FT ACT SITE 269 269 GATASE (BY SIMILARITY).
FT STRAND 4 9
FT TURN 10 11
FT STRAND 14 19
FT STRAND 24 33
FT HELIX 39 43
FT TURN 44 44
FT HELIX 46 48


```

RX MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamachew J.,
RA Gill J., Scariato V., Massignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MCS8."
RL Science 287:1809-1815 (2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE02403; AAF40903.1; -
CC PIR; G81195; G81195.
CC HSSP; P21889; 1EQR.
CC TIGR; NMB0466; -.
CC HAMAP; MF_00044; -.
CC InterPro; IPR004524; Aaps_bact.
CC InterPro; IPR004115; GAD_dom.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR004364; tRNA-synt_2.
CC InterPro; IPR002312; tRNA-synt_asep.
CC InterPro; IPR004365; tRNA_anti.
CC InterPro; IPR006195; tRNA_ligase_II.
CC Pfam; PF02938; GAD; 1.
CC Pfam; PF00152; tRNA-synt_2; 2.
CC Pfam; PF01336; tRNA_anti; 1.
CC PRINTS; PR01042; TRNASYNTHASP.
CC TIGRFAMs; TIGR00459; aaps_bact; 1.
CC PROSITE; PS00862; AA TRNA_LIGASE II; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC KW
CC SEQUENCE 602 AA; 68124 MW; 5DC8A016B0C13E3C CRC64;
CC
CC Query Match 35.0%; Score 52.5; DB 1; Length 602;
CC Best Local Similarity 62.5%; Pred. No. 4.3;
CC Matches 10; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
CC
CC Qy 4 ASSYDYIL-GWEPFGG 18
CC Db 480 ARAYDMVINGWEIGGG 495
CC
CC RESULT 7
CC SYD_STAAM
CC ID -SYD_STAAM STANDARD; PRT; 588 AA.
CC AC Q9YTL9;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
CC (ApsRS).
CC OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
CC OS Staphylococcus aureus (strain N315), and
CC Staphylococcus aureus (strain MW2).
CC OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
CC NCBI_TaxID=158878, 158879, 196620;
CC [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=MU50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Negai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Masumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita N., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiratake K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240 (2001).
CC
CC SEQUENCE FROM N.A.
CC SEQUENCE FROM N.A.
CC SEQUENCE FROM N.A.
CC SEQUENCE FROM N.A.
CC MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano M., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827 (2002).
CC -1- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003363; BAB57792.1; -
CC DR EMBL; AP003134; BAB42722.1; -
CC DR EMBL; AP004827; BAB95445.1; -
CC DR PIR; E89945; E89945.
CC DR HSSP; P21889; 1COA.
CC HAMAP; MF_00044; -.
CC InterPro; IPR004524; Aaps_bact.
CC InterPro; IPR004115; GAD_dom.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR004364; tRNA-synt_2.
CC InterPro; IPR002312; tRNA-synt_asep.
CC InterPro; IPR004365; tRNA_anti.
CC InterPro; IPR006195; tRNA_ligase_II.
CC Pfam; PF02938; GAD; 1.
CC Pfam; PF00152; tRNA-synt_2; 2.
CC Pfam; PF01336; tRNA_anti; 1.
CC PRINTS; PR01042; TRNASYNTHASP.
CC TIGRFAMs; TIGR00459; aaps_bact; 1.
CC PROSITE; PS00862; AA TRNA_LIGASE II; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC KW
CC SEQUENCE 588 AA; 66598 MW; 38A6F2C4BF9E9B9B CRC64;
CC
CC Query Match 33.7%; Score 50.5; DB 1; Length 588;
CC Best Local Similarity 58.8%; Pred. No. 8.3;
CC Matches 10; Conservative 4; Mismatches 2; Indels 1; Gaps 1;
CC
CC Qy 3 EASSYDYIL-GWEPFGG 18
CC Db 473 EARNAYDIVLNGYELGGG 489
CC
CC RESULT 8
CC SYD_HELPY
CC ID -SYD_HELPY STANDARD; PRT; 577 AA.
CC AC P56459;

```

```

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPs OR HP0617.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Keriavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AE000576; AAD07682.1; -.
DR PIR; A64597; A64597.
DR HSSP; P36419; LEFW.
DR TIGR; HP0617; -.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002332; tRNA-synt_2.
DR InterPro; IPR004365; tRNA-anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA-anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; asps_bact; 1.
DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 577 AA; 65601 MW; 29963C376EE9E68F CRC64;

Query Match 33.3%; Score 50; DB 1; Length 577;
Best Local Similarity 52.0%; Pred. No. 9.7;
Matches 13; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 3 EASSYDYL-GWEFFGG-VPEHKKE 25
Db 462 EAHADVVLNGVELGGGSIIRHKEE 486

RESULT 9
SYD_HELPJ

DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
DE (AspRS).
GN ASPs OR JHP0560.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=99120557; PubMed=9923682;
RX Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AE001489; AAD06143.1; -.
DR PIR; C71916; C71916.
DR HSSP; P36419; LEFW.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002332; tRNA-synt_2.
DR InterPro; IPR004365; tRNA-anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA-anti; 1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; asps_bact; 1.
DR PROSITE; PS50862; AA_TRNA_LIGASE_II; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 579 AA; 65616 MW; 4977D97A9562250A CRC64;

Query Match 33.3%; Score 50; DB 1; Length 579;
Best Local Similarity 52.0%; Pred. No. 9.7;
Matches 13; Conservative 4; Mismatches 6; Indels 2; Gaps 2;

QY 3 EASSYDYL-GWEFFGG-VPEHKKE 25
Db 462 EAHADVVLNGVELGGGSIIRHKEE 486

RESULT 10
SYD_LACPJ
ID SYD_LACPJ STANDARD; PRT; 598 AA.
AC Q88VQ8;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```


15-MAR-2004 (Rel. 43, Last annotation update)
 Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase) (AsPRS1).
 GN ASPS OR LP 1980.
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 CC Lactobacillus.
 OX NCBI_TaxID=1590;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIMB 8826 / WCFS1;
 RX MEDLINE=22480296; PubMed=12566566;
 RA Klerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,
 RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
 RA De Vos W.M., Siezen R.J.;
 RT "Complete genome sequence of Lactobacillus plantarum WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).
 CC -|- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -|- SUBUNIT: Homodimer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -|- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL935257; CAD64363.1; -
 CC HAMAP; MF 00044; -; 1.
 CC InterPro; IPR004115; GAD dom.
 CC InterPro; IPR008994; Nucleic_acid_OB.
 CC InterPro; IPR004364; tRNA-synt 2.
 CC InterPro; IPR002312; tRNA-synt asp.
 CC InterPro; IPR004365; tRNA anti-
 CC InterPro; IPR006195; tRNA_ligase_II.
 CC Pfam; PF02938; GAD; 1.
 CC Pfam; PF00152; tRNA-synt 2; 1.
 CC Pfam; PF01336; tRNA anti; 1.
 CC PRINTS; PR01042; TRNASYNTHASP.
 CC PROSITE; PS50862; AA TRNA_LIGASE II; 1.
 CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 CC Complete proteome.
 CC SEQUENCE 598 AA; 67603 MW; 971B6CDBD645DD6D CRC64;
 CC
 CC Query Match 33.3%; Score 50; DB 1; Length 598;
 CC Best Local Similarity 54.2%; Pred. No. 10;
 CC Matches 13; Conservative 3; Mismatches 6; Indels 2; Gaps 2;
 CC
 CC Qy 4 ASSYDYIL-GWFGGG-VPEHKE 25
 CC ||||| : : : : :
 CC 475 AQSVDIILGVLGGGSIHRTRE 498
 CC
 CC RESULT 11
 CC HEMZ BACHD
 CC ID HEMZ BACHD STANDARD; PRT; 310 AA.
 CC AC Q9KDK9;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Ferrochelatase (EC 4.99.1.1) (Protoheme ferro-lyase) (Heme
 CC synthetase).
 CC GN HEMH OR BH1203.
 CC OS Bacillus halodurans.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CC OX NCBI_TaxID=86665;
 CC RN (1)

SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -|- FUNCTION: Catalyzes the ferrous insertion into protoporphyrin IX.
 CC -|- CATALYTIC ACTIVITY: Protoporphyrin + Fe(2+) = protoheme + 2 H(+).
 CC -|- PATHWAY: Protoheme biosynthesis; last step.
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -|- SIMILARITY: Belongs to the ferrochelatase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP001511; BAB04922.1; -
 CC PIN; C83800; C83800.
 CC DR HSSP; P32396; 1C1H.
 CC DR HAMAP; MF 00323; -; 1.
 CC InterPro; IPR001015; Ferrochelatase.
 CC Pfam; PF00762; Ferrochelatase; 1.
 CC ProDom; PD002792; Ferrochelatase; 1.
 CC TRGFAMS; TIGR00109; hemH; 1.
 CC PROSITE; PS00534; FERROCHELATASE; 1.
 CC PROSITE; PS00534; FERROCHELATASE; 1.
 CC Porphyrin biosynthesis; Heme biosynthesis; Lyase; Iron;
 CC Complete proteome.
 CC METAL 181 181 IRON (BY SIMILARITY).
 CC FT METAL 262 262 IRON (BY SIMILARITY).
 CC FT METAL 262 262 IRON (BY SIMILARITY).
 CC SQ SEQUENCE 310 AA; 35271 MW; AA28CE33BAB1FF21 CRC64;
 CC
 CC Query Match 32.0%; Score 48; DB 1; Length 310;
 CC Best Local Similarity 42.9%; Pred. No. 10;
 CC Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
 CC
 CC Qy 1 AKEASSYDYILGWFGGGVPE 21
 CC ||||| : : : : :
 CC 206 AKGAGITNVAVGWQSGNTPE 226
 CC
 CC RESULT 12
 CC SYD STAEF
 CC ID SYD STAEF STANDARD; PRT; 588 AA.
 CC AC Q8CS99;
 CC DT 15-MAR-2004 (Rel. 43, Created)
 CC DT 15-MAR-2004 (Rel. 43, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 CC (AsPRS).
 CC GN ASPS OR SE1111.
 CC OS Staphylococcus epidermidis.
 CC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 CC OX NCBI_TaxID=1282;
 CC RN (1)
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=ATCC 12228;
 CC RX PubMed=12950922;
 CC RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
 CC Qiu Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
 CC Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;
 CC "Genome-based analysis of virulence genes in a non-biofilm-forming
 CC Staphylococcus epidermidis strain (ATCC 12228).";
 CC Mol. Microbiol. 49:1577-1593(2003).
 CC -|- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -|- SUBUNIT: Homodimer (By similarity).

RT is essential for Wnt-3a-induced accumulation of beta-catenin. ";
 RL J. Biol. Chem. 278:14066-14073 (2003).
 CC -!- FUNCTION: Casein kinases are operationally defined by their
 CC preferential utilization of acidic proteins such as caseins
 CC as substrates. It can phosphorylate a large number of proteins.
 CC Participates in Wnt signaling. Phosphorylates DVL1.
 CC -!- SUBUNIT: Monomer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- PTM: Autophosphorylated (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SRP/THR FAMILY OF PROTEIN KINASES.
 CC CASEIN KINASE I SUBFAMILY. COULD BE THE ORTHOLOG OF DROSOPHILA
 CC DOUBLE-TIME.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/announcement/>
 CC or send an email to license@ebi-sib.ch).
 CC
 CC EMBL; L37043; AAC41761.1; -
 CC EMBL; AB024597; BAA92345.1; -
 CC EMBL; AB091043; BAC10902.1; -
 CC EMBL; AL020993; CAAL5888.1; -
 CC EMBL; BC006490; AAH06490.1; -
 CC PIR; I61744; I61744.
 CC HSP; Q06486; 1CKJ.
 CC Genew; HGNC:2453; CSNK1E.
 CC MIM; 600863; -
 CC GO; GO:0004681; P:casein kinase I activity; TAS.
 CC GO; GO:0006281; P:DNA repair; TAS.
 CC GO; GO:0006468; P:protein amino acid phosphorylation; TAS.
 CC GO; GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR008271; Ser_Thr_kin_AS.
 CC Pfam; PF00069; pkinase; 1.
 CC ProDom; PD00001; Prot_kinase; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
 CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding;
 CC phosphorylation; Multigene family.
 CC DOMAIN 9 277 PROTEIN KINASE.
 CC NP BIND 15 23 ATP (BY SIMILARITY).
 CC BINDING 38 38 ATP (BY SIMILARITY).
 CC ACT SITE 128 128 BY SIMILARITY.
 CC SEQUENCE 416 AA; 47315 MW; EBI698AE914324 CRC64;
 CC
 CC Query Match 31.3%; Score 47; DB 1; Length 416;
 CC Best Local Similarity 29.7%; Pred. No. 19;
 CC Matches 11; Conservative 6; Mismatches 8; Indels 12; Gaps 2;
 CC
 CC QY 2 KEASDYIILGW---EFGG-----VPEHKKE 26
 CC :: ||||| : : : : :
 CC Db 279 RQGSYDYVDFWMLKFGAARNPVDVRRERHEREE 315
 CC :: ||||| : : : : :
 CC
 CC RESULT 14
 CC SYD_LISIN
 CC ID -SYD_LISIN STANDARD; PRT; 591 AA.
 CC AC Q92BJ4;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 CC DE (AspRS).
 CC GN ASPS OR L1N1554.
 CC OS Listeria innocua.
 CC OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 CC NCBI_TaxID=1642;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.

RC STRAIN=CLIP 11262 / Serovar 6a;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
 RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 RA Jones L.-M., Kaerst U., Kref J., Kuhn M., Kunst F., Kurapkat G.,
 RA Madueno E., Maitournam A., Mata Vicente J., NG E., Nedjari H.,
 RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez P.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RL "Comparative genomics of *Listeria species*."; Science 294:849-852(2001).
 CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
 CC diphosphate + L-aspartyl-tRNA(Asp).
 CC -!- SUBUNIT: Homodimer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis-sib.ch/announcement/>
 CC or send an email to license@ebi-sib.ch).
 CC
 CC EMBL; AL596169; CAC96785.1; -
 CC PIR; A11626; A11626.
 CC ListList; LIN01554; -
 CC HAMAP; MF_00044; -; 1.
 CC InterPro; IPR004524; Asps_bact.
 CC InterPro; IPR004115; GAD_dom.
 CC InterPro; IPR008994; Nucleic_acid_OB.
 CC InterPro; IPR004364; tRNA-synt_2.
 CC InterPro; IPR002312; tRNA-synt_2.
 CC InterPro; IPR004365; tRNA-anti-asp.
 CC InterPro; IPR006195; tRNA_ligase_II.
 CC Pfam; PF02938; GAD; 1.
 CC Pfam; PF00152; tRNA-synt_2; 2.
 CC Pfam; PF01336; tRNA_anti; 1.
 CC PRINTS; PR01042; TRNASYNTHASP.
 CC TIGRFAMs; TIGR00459; asps_bact; 1.
 CC PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
 CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 CC Complete proteome.
 CC SEQUENCE 591 AA; 66374 MW; 81ABAC0D02147F7E CRC64;
 CC
 CC Query Match 31.3%; Score 47; DB 1; Length 591;
 CC Best Local Similarity 50.0%; Pred. No. 28;
 CC Matches 12; Conservative 5; Mismatches 5; Indels 2; Gaps 2;
 CC
 CC QY 4 ASSDYVIL-GWEPGGG-VPEHKKE 25
 CC :: ||| : : : : :
 CC Db 473 AEAYDIVLNGYEIGGSLRIYKKE 496
 CC :: ||| : : : : :
 CC
 CC RESULT 15
 CC SYD_LISMO
 CC ID -SYD_LISMO STANDARD; PRT; 591 AA.
 CC AC Q8Y709;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)
 CC DE (AspRS).
 CC GN ASPS OR LMO1519.
 CC OS Listeria monocytogenes.
 CC OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
 CC NCBI_TaxID=1639;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.

```

RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a:
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rueniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Gobel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournan A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Roland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP +
CC diphosphate + L-aspartyl-tRNA(Asp).
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL591979; CAC959597.1; -.
DR PIR; AG1264; AG1264.
DR List18c; LM001519; -.
DR HAMAP; MF_00044; -.
DR InterPro; IPR004524; Asps_bact.
DR InterPro; IPR004115; GAD_dom.
DR InterPro; IPR008994; Nucleic_acid_OB.
DR InterPro; IPR004364; tRNA-synt_2.
DR InterPro; IPR002312; tRNA-synt_2.
DR InterPro; IPR004365; tRNA_anti.
DR InterPro; IPR006195; tRNA_ligase_II.
DR Pfam; PF02938; GAD; 1.
DR Pfam; PF00152; tRNA-synt_2; 2.
DR Pfam; PF01336; tRNA_anti_1.
DR PRINTS; PR01042; TRNASYNTHASP.
DR TIGRFAMs; TIGR00459; asps_bact; 1.
DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
DR Complete proteome.
SQ SEQUENCE 591 AA; 66382 MW; DE143FD009106BBA CRC64;

```

```

Query Match      31.3%; Score 47; DB 1; Length 591;
Best Local Similarity 50.0%; Pred.No. 28;
Matches 12; Conservative 5; Mismatches 5; Indels 2; Gaps 2;

```

```

QY      4 ASSYDYIL-GWERGGG-VPEHKKE 25
DB      473 AEAYDIVLNGYEIGGGSLRIYKKE 496

```

Search completed: September 22, 2004, 18:06:59
Job time : 7.57851 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:32 ; Search time 5.16529 Seconds
(without alignments)
252.020 Million cell updates/sec

Title: US-10-774-602-11
Perfect score: 124
Sequence: 1 YEKANAYOKANQAVLKAEASSYD 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	53	42.7	463	1 K10_DROME	P13468 drosophila
2	50	40.3	106	1 GLPE_VIBCH	Q9KVP1 vibrio chol
3	48	38.7	988	1 TNPS_PSEAE	P06695 pseudomonas
4	47.5	38.3	100	1 RL23_ECOLI	P02424 escherichia
5	47.5	38.3	472	1 UXAC_OCEIH	O8CXN4 oceanobacil
6	47.5	38.3	1131	1 RFL1_MOUSE	P13501 mus musculu
7	47	37.9	51	1 Y070_RICCN	Q92JJ7 rickettsia
8	47	37.9	119	1 RL20_NEIMA	Q9JVA1 neisseria m
9	47	37.9	119	1 RL20_NEIMB	Q9K093 neisseria m
10	47	37.9	541	1 CAT4_ASCSU	P90682 ascaris suu
11	47	37.9	567	1 PGTA_MOUSE	Q9JHK4 mus musculu
12	46	37.1	119	1 RL20_CHRVO	O7NVY3 chromobacte
13	46	37.1	119	1 RL20_NITEU	O82VY4 nitrosomona
14	46	37.1	338	1 YQJM_BACSU	P54550 bacillus su
15	46	37.1	503	1 AMPA_YERPE	Q8ZBH3 yersinia pe
16	45	36.3	124	1 RS16_RHIME	Q92143 rhizobium m
17	45	36.3	431	1 IM44_YEAST	O01852 saccharomyc
18	45	36.3	632	1 NTP1_SPVKA	Q9Q822 shope fibro
19	45	36.3	1086	1 RNC_CABEL	O01326 caenorhabdi
20	45	36.3	1189	1 SMC2_CHICK	Q90988 gallus gall
21	44.5	35.9	273	1 NADC_HELPY	O25909 helicobacte
22	44.5	35.9	503	1 AMPA_ECOLI	P11648 escherichia
23	44.5	35.9	503	1 AMPA_SALTY	O81116 salmonella
24	44.5	35.9	503	1 AMPA_SALTY	O82K29 salmonella
25	44.5	35.9	693	1 YD01_SCHPO	O14286 schistosom
26	44	35.5	103	1 GLPE_HAEIN	P44819 haemophilus
27	44	35.5	155	1 LEA2_CICAR	O49817 cicor ariet
28	44	35.5	177	1 LEA1_CICAR	O49816 cicor ariet
29	44	35.5	199	1 WRBA_BRAJA	O89D74 bradyrhizob
30	44	35.5	267	1 IF2A_ARCFU	O29723 archaeoglob
31	44	35.5	231	1 SNAG_ARATH	Q9SP65 arabidopsis
32	44	35.5	428	1 PROA_TREPA	P74935 treponema p
33	44	35.5	534	1 HTR2_NATPH	P42259 natronomona

RESULT 1

ID	K10_DROME	STANDARD;	PRT;	463 AA.
AC	P13468; O46075; Q9W505;			
DT	01-JAN-1990 (Rel. 13, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	DNA-binding protein K10 (female sterile protein K10).			
GN	FS(1)K10 OR EG:30B8.5 OR CG3218.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Sphyroidea; Drosophilidae; Drosophila.			
ON	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Prost E., Deryckere F., Roos C., Haenlin M., Pantescio V.,			
RA	Mohier V.;			
RT	"Role of the oocyte nucleus in determination of the dorsoventral			
RT	polarity of Drosophila as revealed by molecular analysis of the K10			
RT	gene.";			
RL	Genes Dev. 2:891-900(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Oregon-R;			
RX	MEDLINE=20196011; PubMed=10731137;			
RA	Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,			
RA	Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu B.,			
RA	Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,			
RA	Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,			
RA	Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablo B.,			
RA	Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F.,			
RA	Beinert N., Dove G., Schaefer U., Jaekle H., Bucheton A.,			
RA	Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,			
RA	McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,			
RA	Glover D.M.;			
RT	"From sequence to chromosome: the tip of the X chromosome of D.			
RT	melanogaster.";			
RL	Science 287:2220-2222(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Berkely;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,			
RA	Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,			
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			

Q2696 homo sapien
Q08602 rattus norv
P09032 saccharomyc
Q07093 drosophila
P31384 saccharomyc
P33438 drosophila
P37267 saccharomyc
P43814 haemophilus
P17486 caenorhabdi
P45730 populus tri
Q04593 pisum sativ
Q8ey85 leptospira

ALIGNMENTS

[illegible]

"Structure of the Escherichia coli S10 ribosomal protein operon."; Nucleic Acids Res. 13:4521-4526(1985).

[2]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=K12 / MG1655;
MEDLINE=97426617; PubMed=9278503;
Blaattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
"The complete genome sequence of Escherichia coli K-12."; Science 277:1453-1474(1997).

[3]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
Weil J., Goldberg M.B., Burland V., Plunkett G. III, Redford P., Roesch P., Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D., Mayhew H.L.T., Donnenberg M.S., Blattner F.R.;
"Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli."; Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).

[4]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
Weil J., Goldberg M.B., Burland V., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Podocaj J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."; Nature 409:529-533(2001).

[5]
SEQUENCE FROM N.A.
SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
Weil J., Goldberg M.B., Burland V., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Podocaj J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;
"Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."; Nature 409:529-533(2001).

[6]
SEQUENCE.
SPECIES=E.coli; STRAIN=K12;
MEDLINE=80092111; PubMed=391594;
Wittmann-Liebold B., Greuer B.;
Primary structure of protein L23 from the Escherichia coli serotype 2a;
FEBS Lett. 108:69-74 (1979).

[7]
SEQUENCE FROM N.A.
SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
MEDLINE=2272406; PubMed=12384590;
Lin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H., Yang J., Yang F., Zhang J., Yang G., Wu H., Qu D., Dong J., J. L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S., Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou S., J.;
"Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157."; Nucleic Acids Res. 30:4432-4441(2002).

[8]
SEQUENCE FROM N.A.
SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
MEDLINE=22590274; PubMed=12704152;
Weil J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W., Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A., Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S., J.;

RT "Cloning, expression, and chromosomal localization of the
RT 140-kilodalton subunit of replication factor C from mice and
RT humans.";
RL Mol. Cell. Biol. 14:1626-1634 (1994).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=Swiss;
RX MEDLINE=95388065; PubMed=7659092;
RA McGehee Habener J.F.;
RT "Differentiation-specific element binding protein (DSEB) binds to a
RT defined element in the promoter of the angiotensinogen gene required
RT for the irreversible induction of gene expression during
RT differentiation of 3T3-L1 adipoblasts to adipocytes.";
RL Mol. Endocrinol. 9:487-501 (1995).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RA Haque S.J.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-565 FROM N.A.
RC STRAIN=LAF1;
RX Haque S.J., Kumar A., Fischer T., Rutherford M.N., Williams B.R.;
RA "Evaluation of inter- and intramolecular primary structure homologies
RT of interferons by a Monte Carlo method.";
RL J. Interferon Res. 10:31-31 (1990).
CC -1- FUNCTION: THE ELONGATION OF PRIMED DNA TEMPLATES BY DNA POLYMERASE
CC DELTA AND EPSILON REQUIRES THE ACTION OF THE ACCESSORY PROTEINS
CC PCNA AND ACTIVATOR 1. THE 140 SUBUNIT BINDS TO THE PRIMER-
CC TEMPLATE JUNCTION.
CC -1- SUBUNIT: HETEROPEPTAMER OF SUBUNITS OF 140/145, 40, 38, 37, AND
CC 36.5 kDa THAT FORMS A COMPLEX WITH PCNA IN THE PRESENCE OF ATP.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P35601-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P35601-2; Sequence=VSP_008444;
CC Note=Alternative use of an acceptor site;
CC -1- SIMILARITY: Belongs to the activator 1 140 kDa subunit family.
CC -1- SIMILARITY: Contains 1 BRCT domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U01222; AAA21643.1; -
CC EMBL; X72711; CAA51260.1; -
CC EMBL; U36441; AAA75698.1; -
CC EMBL; U07157; AAC52140.1; -
CC EMBL; U15037; AAB60452.1; -
CC PIR; A49393; A49393.
CC MGD; MGI:97891; Reccl1.
CC GO; GO:0005634; C:nucleus; IDA.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003959; AAA_ATPase_central.
CC InterPro; IPR001357; BRCT.
CC Pfam; PF00004; AAA; 1.
CC Pfam; PF00533; BRCT; 1.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00292; BRCT; 1.
CC PROSITE; PSS0172; BRCT; 1.
KW DNA replication; ATP-binding; Transcription regulation; DNA-binding;

KW Activator; Nuclear protein; Zinc-finger; Alternative splicing.
FT DOMAIN 399 477 BRCT.
FT NP_BIND 635 642 ATP (BY SIMILARITY).
FT ZN_FING 734 751 C2HC-TYPE (POTENTIAL).
FT DOMAIN 1104 1108 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 354 528 INTERFERON-STIMULATED-RESPONSE-ELEMENT
FT BINDING REGION.
FT VARSPLIC 614 614 Missing (in isoform 2).
FT /FTID=VSP_008444.
FT CONFLICT 66 66 Y -> N (IN REF. 3).
FT CONFLICT 187 187 E -> EPDFCLSLIPGQI (IN REF. 4).
FT CONFLICT 254 254 V -> A (IN REF. 5).
FT CONFLICT 559 559 N -> S (IN REF. 4).
FT CONFLICT 945 945 S -> N (IN REF. 1).
FT CONFLICT 1071 1071 T -> A (IN REF. 3).
FT CONFLICT 1104 1104 K -> Q (IN REF. 4).
SQ SEQUENCE 1131 AA; 125984 MW; ASF4F970A7F9EE94 CRC64;
Query Match 38.3%; Score 47.5; DB 1; Length 1131;
Best Local Similarity 36.4%; Pred. No. 67;
Matches 12; Conservative 7; Mismatches 5; Indels 9; Gaps 1;
QY 2 EKAKNAYQK-----ANQAVLKAKEASSYD 25
DB 296 QPKSAHRKEACSSPKASAKALMKAKESSYN 328
RESULT 7
Y070_RICCN STANDARD; PRT; 51 AA.
ID Y070_RICCN
AC Q92JJ7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein RC0070.
GN RC0070.
OS Rickettsia conorii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Malish 7;
RX MEDLINE=21442074; PubMed=11557893;
RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA Raoult D.;
RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii";
RL Science 293:2093-2098 (2001).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB008574; AAL02608.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 51 AA; 5983 MW; 928D5580361EA138 CRC64;
Query Match 37.9%; Score 47; DB 1; Length 51;
Best Local Similarity 45.0%; Pred. No. 3.9;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 1 YEKAKNAYQKANKAVLKAKE 20
DB 10 YKGNRYKKKEFEMLKCEE 29
RESULT 8
RL20_NEIMA

```

ID  RL20_NEIMA  STANDARD;  PRT;  119 AA.
AC  Q9JVA1;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L20.
GN  RPLT OR NMA0932.
OS  Neisseria meningitidis (serogroup A).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC  Neisseriaceae; Neisseria.
OX  NCBI_TaxID=65699;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=22491 / Serogroup A / Serotype 4A;
RX  MEDLINE=20222556; PubMed=10761919;
RA  Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA  Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA  Davies R.M., Davis P., Feltwell T., Hamlin N., Holroyd S.,
RA  Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA  Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA  Whitehead S., Spratt B.G., Barrall B.G.;
RT  "Complete DNA sequence of a serogroup A strain of Neisseria
RT  meningitidis 22491."
RL  Nature 404:502-506(2000).
CC  -!- FUNCTION: This protein binds directly to 23s ribosomal RNA and is
CC  necessary for the in vitro assembly process of the 50s ribosomal
CC  subunit. It is not involved in the protein synthesizing functions
CC  of that subunit (By similarity).
CC  -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AL162754; CAB84204.1; -.
CC  PIR; E81939; E81939.
CC  HAMAP; MF_00382; -.
CC  InterPro; IPR005813; Ribosomal_L20.
CC  Pfam; PF00453; Ribosomal_L20; 1.
CC  PRINTS; PR00062; RIBOSOMALL20.
CC  ProDom; PD002389; L20; 1.
CC  TIGRFAMs; TIGR01032; rplT_bact; 1.
CC  PROSITE; PS00937; RIBOSOMAL_L20; 1.
CC  Ribosomal protein; rRNA-binding; Complete proteome.
SQ  SEQUENCE 119 AA; 13680 MW; 9B9859BB5550D650 CRC64;

Query Match 37.9%; Score 47; DB 1; Length 119;
Best Local Similarity 43.5%; Pred. No. 8.8;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY  3 KAKNAYQKANAQVLKAKEASSYD 25
Db  27 RKKNVTRVAKQAVMKAGQYAYRD 49

RESULT 9
RL20_NEIMB
ID  RL20_NEIMB  STANDARD;  PRT;  119 AA.
AC  Q9K093;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  50S ribosomal protein L20.
GN  RPLT OR NMB0723.
OS  Neisseria meningitidis (serogroup B).
OC  Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC  Neisseriaceae; Neisseria.
OX  NCBI_TaxID=491;

```

```

RN  SEQUENCE FROM N.A.
RP  STRAIN=MC58 / Serogroup B;
RX  MEDLINE=20175755; PubMed=10710307;
RA  Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA  Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA  Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA  Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA  Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B.,
RA  Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA  Gill J., Scariato V., Masiagnani V., Pizzo M., Grandi G., Sun L.,
RA  Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT  "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT  MC58."
RL  Science 287:1809-1815(2000).
CC  -!- FUNCTION: This protein binds directly to 23s ribosomal RNA and is
CC  necessary for the in vitro assembly process of the 50s ribosomal
CC  subunit. It is not involved in the protein synthesizing functions
CC  of that subunit (By similarity).
CC  -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; AF002427; AAF41136.1; -.
CC  PIR; F81165; F81165.
CC  TIGR; NMB0723; -.
CC  HAMAP; MF_00382; -.
CC  InterPro; IPR005813; Ribosomal_L20.
CC  Pfam; PF00453; Ribosomal_L20; 1.
CC  PRINTS; PR00062; RIBOSOMALL20.
CC  ProDom; PD002389; L20; 1.
CC  TIGRFAMs; TIGR01032; rplT_bact; 1.
CC  PROSITE; PS00937; RIBOSOMAL_L20; 1.
CC  Ribosomal protein; rRNA-binding; Complete proteome.
SQ  SEQUENCE 119 AA; 13710 MW; 9B98430B554A6650 CRC64;

Query Match 37.9%; Score 47; DB 1; Length 119;
Best Local Similarity 43.5%; Pred. No. 8.8;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY  3 KAKNAYQKANAQVLKAKEASSYD 25
Db  27 RKKNVTRVAKQAVMKAGQYAYRD 49

RESULT 10
CATA_ASCSU
ID  CATA_ASCSU  STANDARD;  PRT;  541 AA.
AC  P90682;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Catalase (EC 1.11.1.6).
GN  CAT.
OS  Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC  Ascarididae; Ascaris.
OX  NCBI_TaxID=6253;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Eckelt V.H.O.;
RL  Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: Occurs in almost all aerobically respiring organisms and
CC  serves to protect cells from the toxic effects of hydrogen
CC  peroxide.
CC  -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.

```

CC -1- COFACTOR: Heme group (By similarity).
 CC -1- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
 CC -1- SIMILARITY: Belongs to the catalase family.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; Y10611; CAA71618.1; -;
 DR HSP; P04040; IP4J.
 DR InterPro; IPR002226; Catalase.
 DR Pfam; PF00199; Catalase; 1.
 DR PRINTS; PR00067; CATALASE.
 DR PRODOM; PD000510; Catalase; 1.
 DR PROSITE; PS00437; CATALASE_1; 1.
 DR PROSITE; PS00438; CATALASE_2; 1.
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
 K Peroxisome.
 FT ACT_SITE 74 74 BY SIMILARITY.
 FT ACT_SITE 147 147 BY SIMILARITY.
 FT METAL 357 357 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 541 AA; 61962 MW; 70890E1AECA170B0 CRC64;
 Query Match 37.9%; Score 47; DB 1; Length 541;
 Best Local Similarity 56.2%; Pred. No. 38;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 Oy 5 KNAYQKQNAVILKAKE 20
 Db 493 RNALQKANEAMQKKE 508
 :|||:||||:|:|:
 RESULT 11
 PGTA_MOUSE STANDARD; PRT; 567 AA.
 AC Q9JHK4; Q9JLK2;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Geranylgeranyl transferase type II alpha subunit (EC 2.5.1.60) (Rab
 DE geranylgeranyl transferase alpha subunit) (Rab geranyl-
 DE geranyltransferase alpha subunit) (Rab GG transferase alpha) (Rab
 DE GGTase alpha).
 GN RABGGTA.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND DISEASE.
 RC STRAIN=C57BL/6J-GM/Gm, and C57BL/6J;
 RX MEDLINE=20226072; PubMed=10737774;
 RA Dettler J.C., Zhang Q., Mules E.H., Novack E.K., Mishra V.S., Li W.,
 RA McMurtrie E.B., Tchernev V.T., Wallace M.R., Seabra M.C., Swank R.T.,
 RA Kingmore S.K.;
 RA "Rab geranylgeranyl transferase alpha mutation in the gunmetal mouse
 RT reduces Rab prenylation and platelet synthesis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:4144-4149(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6J, TISSUE=Kidney;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: Catalyzes the transfer of a geranyl-geranyl moiety from
 CC geranyl-geranyl pyrophosphate to both cysteines in Rab proteins
 CC with an -XXCC, -XCXC and -CCXX C-terminal, such as RAB1A, RAB3A
 CC and RAB5A respectively (By similarity).
 CC -1- CATALYTIC ACTIVITY: 2 geranylgeranyl diphosphate + protein-
 CC cysteine = 2 S-geranylgeranyl-protein + 2 diphosphate.
 CC -1- ENZYME REGULATION: The enzymatic reaction requires the aid of a
 CC Rab escort protein (also called component A) (By similarity).
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit, collectively
 CC called component B (By similarity).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q9JHK4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9JHK4-2; Sequence=VSP_009113, VSP_009114;
 CC -1- DISEASE: Defects in RABGGTA are the cause of the gunmetal (gm)
 CC phenotype. Mice homozygous for gm have prolonged bleeding,
 CC thrombocytopenia and reduced platelet alpha- and delta-granula
 CC contents.
 CC -1- SIMILARITY: Belongs to the protein prenyltransferase alpha subunit
 CC family.
 CC -1- SIMILARITY: Contains 5 PPTA repeats.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration -
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR ENBL; AF127654; AAF65918.1; -;
 DR ENBL; AF127655; AAF65919.1; -;
 DR ENBL; AF127656; AAF65920.1; -;
 DR ENBL; AF127658; AAF65921.1; -;
 DR ENBL; AF127659; AAF65922.1; -;
 DR ENBL; AF127660; AAF65923.1; -;
 DR ENBL; AF127662; AAF65924.1; -;
 DR ENBL; AK008625; BAB22240.1; -;
 DR MGD; MGI:1860443; Rabggta.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007092; LRR SDS322.
 DR InterPro; IPR002088; PPTA.
 DR InterPro; IPR008940; Prenyl trans.
 DR InterPro; IPR009087; RabGGT_A_insert.
 DR Pfam; PF00560; LRR; 3.
 DR Pfam; PF01239; PPTA; 5.
 DR PRINTS; PR00019; LEURICHRPT.
 KW Transferase; Prenyltransferase; Repeat; Alternative splicing.
 FT REPEAT 47 80 PPTA 1.
 FT REPEAT 91 124 PPTA 2.
 FT REPEAT 127 160 PPTA 3.
 FT REPEAT 162 195 PPTA 4.
 FT REPEAT 210 243 PPTA 5.
 FT VARSPPLIC 337 339 HOE -> DAV (in isoform 2).
 FT VARSPPLIC 340 567 /FTId=VSP_009113.
 FT VARSPPLIC 340 567 Missing (in isoform 2).
 SQ SEQUENCE 567 AA; 64989 MW; 387DA2DAC12C4C0D CRC64;

Query Match 37.9%; Score 47; DB 1; Length 567;
Best Local Similarity 41.7%; Pred. No. 40;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 EKAKNAYQKANOAVLKAKEASSYD 25
: ||| ||| ||| : ||| :
Db 22 EQKLIKYSATQAVQFKREAGELD 45

RESULT 12
RL20 CHRVO STANDARD; PRT; 119 AA.
AC Q7NYC3;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein L20.
GN RPLT OR CVI351.
OC Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=2282880; PubMed=14500782;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,
Batista-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
Bordignon J., Brigidio M.M., Brito C.A., Brocchi M., Burity H.A.,
Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
Fantinatti P., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
Grattapaglia D., Griseard E.C., Hanna E.S., Jardim S.N., Laurino J.,
Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
Madeira H.M.F., Manfro G.P., Maranhao A.Q., Martins W.S.,
di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J.J., Pereira J.O.,
Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
Santos E.B.P., Santos F.R., Schneider M.P.C., Seunanez H.N.,
Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
remarkable and exploitable bacterial adaptability";
Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
CC -!- FUNCTION: This protein binds directly to 23S ribosomal RNA and is
necessary for the in vitro assembly process of the 50S ribosomal
subunit. It is not involved in the protein synthesizing functions
of that subunit (By similarity).
CC -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AE016914; AAQ59026.1; -.
DR HAMAP; MF_00382; -; 1.
DR PROSITE; PS00937; RIBOSOMAL_L20; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.

SQ SEQUENCE 119 AA; 13626 MW; 2D74991792880CEA CRC64;
Query Match 37.1%; Score 46; DB 1; Length 119;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 KAKNAYQKANOAVLKAKEASSYD 25
: ||| ||| ||| : ||| :
Db 27 RRKNVYRIAKQAVMKAGQYAYRD 49

RESULT 13
RL20 NITEU STANDARD; PRT; 119 AA.
AC Q82VW4;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 50S ribosomal protein L20.
GN RPLT OR NE0955.
OC Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
Arciero D.M., Hommes N.G., Whitaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
obligate chemolithoautotroph Nitrosomonas europaea";
J. Bacteriol. 185:2759-2773(2003).
RL -!- FUNCTION: This protein binds directly to 23S ribosomal RNA and is
necessary for the in vitro assembly process of the 50S ribosomal
subunit. It is not involved in the protein synthesizing functions
of that subunit (By similarity).
CC -!- SIMILARITY: Belongs to the L20P family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; BX321859; CAD84866.1; -.
DR HAMAP; MF_00382; -; 1.
DR InterPro; IPR005813; Ribosomal_L20.
DR InterPro; IPR005812; Ribosomal_L20b/o.
DR Pfam; PF00453; Ribosomal_L20; 1.
DR PRINTS; PR00062; RIBOSOMALL20.
DR ProDom; PD002389; L20; 1.
DR TIGSFAME; TIGSF01032; rplT bact; 1.
DR PROSITE; PS00937; RIBOSOMAL_L20; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 119 AA; 13724 MW; 779B0490A37D6B4B CRC64;

Query Match 37.1%; Score 46; DB 1; Length 119;
Best Local Similarity 43.5%; Pred. No. 12;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 KAKNAYQKANOAVLKAKEASSYD 25
: ||| ||| ||| : ||| :
Db 27 RRKNVYRIAKQAVMKAGQYAYRD 49

RESULT 14
YQJM_BACSU STANDARD; PRT; 338 AA.
ID YQJM_BACSU
AC P54550;

DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DT Probable NADH-dependent flavin oxidoreductase yqjM (BC 1.-.-.-).
 GN YQM OR BSU23820.
 DE Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168 / JH642;
 RX MEDLINE=97124195; PubMed=8969508;
 RA Mizuno M., Masuda S., Takemaru K.-I., Hosono S., Sato T., Takeuchi M.,
 RA Kobayashi Y.;
 RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of
 RT the Bacillus subtilis genome containing the skin element and many
 RT sporulation genes.";
 RL Microbiology 142:3103-3111(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
 RA Entian K.D., Errington J., Fabret C., Fertari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haehtel J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaubner-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rev M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter P., Scofione F.,
 RA Skiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Takakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipatt A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256(1997).
 CC -1- SIMILARITY: TO THE OYE FAMILY OF NADH-DEPENDENT FLAVIN
 CC OXIDOREDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D84432; BAA12619.1; -;
 CC ENBL; Z99116; CAB14314.1; -;
 CC PIR; E69964; E69964.
 CC Subtilist; EG11742; yqjM.
 CC InterPro; IPR001155; Oxidored FMN.
 CC Pfam; PF00724; oxidored FMN; I.
 CC Hypothetical protein; Oxidoreductase; NAD; FAD; Flavoprotein;
 CC Complete proteome.
 CC SEQUENCE 338 AA; 37583 MW; BDC52D34236326FE CRC64;
 SQ

Query Match 37.1%; Score 46; DB 1; Length 338;
 Best Local Similarity 50.0%; Pred. No. 33;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 2 EKAKNAYQKANOAVLKAKA 21
 DB 137 EKVKTVOEFKQAAAKA 156
 RESULT 15
 AMPA_YERPE
 ID AMPA_YERPE STANDARD; PRT; 503 AA.
 AC Q8ZBH3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
 DE (LAP) (Leucyl aminopeptidase).
 GN PEPA OR YPO3441 OR Y0746.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 CC -1- FUNCTION: Presumably involved in the processing and regular
 CC turnover of intracellular proteins. Catalyzes the removal of
 CC unsubstituted N-terminal amino acids from various peptides (By
 CC similarity).
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
 CC Xbb-, in which Xaa is preferably Leu, but may be other amino acids
 CC including Pro although not Arg or Lys, and Xbb may be Pro.
 CC -1- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to peptidase family M17.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ414157; CAC92671.1; -;
 CC EMBL; AE013676; AAM84333.1; -;
 CC PIR; AC0418; AC0418.
 CC MEROPS; M17.003; -;
 CC HAMAP; MF_00181; -; 1.
 CC InterPro; IPR000819; Peptidase_M17_C.

DR InterPro; IPR008283; Peptidase M17_N.
DR Pfam; PF00983; Peptidase M17; I.
DR Pfam; PF02789; Peptidase M17_N; 1.
DR PRINTS; PR00481; LAMNOPPTDASE.
DR PROSITE; PS00631; CYTOSOL_AP; 1.
KW Hydrolase; Aminopeptidase; Manganese; Complete proteome.
FT ACT_SITE 282 282 POTENTIAL.
FT ACT_SITE 356 356 POTENTIAL.
FT METAL 270 270 MANGANESE 2 (BY SIMILARITY).
FT METAL 275 275 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 293 293 MANGANESE 2 (BY SIMILARITY).
FT METAL 352 352 MANGANESE 1 (BY SIMILARITY).
FT METAL 354 354 MANGANESE 1 AND 2 (BY SIMILARITY).
SQ SEQUENCE 503 AA; 54796 MW; E3A27876304218F2 CRC64;

Query Match 37.1%; Score 46; DB 1; Length 503;
Best Local Similarity 55.6%; Pred. NO. 49;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

3 KAKNAYQKANKQAVLKAKE 20
118 KGRNTYWKVRQAVETAKE 135

Search completed: September 22, 2004, 18:06:57
Job time : 14.1653 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:03:03 ; Search time 12.719 Seconds
(without alignments)
109.592 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKKASSYDYLWFGVGGVPEHKKEEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	100.0	27	3	US-08-416-711-3
2	150	100.0	27	4	US-09-356-497-3
3	150	100.0	64	3	US-08-416-711-1
4	150	100.0	64	4	US-09-356-497-1
5	50.5	33.7	597	4	US-09-252-991A-32073
6	49	32.7	490	4	US-09-232-225-41
7	49	32.7	509	4	US-09-232-225-35
8	49	32.7	509	4	US-09-232-225-38
9	48	32.0	306	4	US-09-386-642-53
10	48	32.0	319	4	US-09-386-642-12
11	48	32.0	560	4	US-09-252-991A-22343
12	48	32.0	671	4	US-09-252-991A-19016
13	47.5	31.7	593	4	US-09-134-001C-3592
14	47	31.3	168	4	US-09-543-681A-6335
15	47	31.3	416	3	US-09-100-664A-9
16	47	31.3	416	4	US-09-335-983-9
17	47	31.3	416	4	US-09-553-867A-9
18	47	31.3	416	4	US-09-553-867A-16
19	47	31.3	416	4	US-09-553-867A-18
20	47	31.3	416	4	US-09-472-112-1
21	47	31.3	416	4	US-09-252-991A-18079
22	47	31.3	947	4	US-09-252-991A-21398
23	47	31.3	1390	2	US-08-770-544-2
24	47	31.3	1390	4	US-09-579-259-2
25	47	31.3	1390	4	US-09-650-324A-2
26	46.5	31.0	523	3	US-08-606-505B-67
27	46.5	31.0	523	3	US-09-616-990-67

Sequence 4, Appli
Sequence 4, Appli
Sequence 4654, Ap
Sequence 797, App
Sequence 21, Appl
Sequence 15, Appl
Sequence 18, Appl
Sequence 6086, Ap
Sequence 4524, Ap
Sequence 101, App
Sequence 84, Appl
Sequence 84, Appl
Sequence 5365, Ap
Sequence 9709, Ap
Sequence 23, Appl
Sequence 7, Appl

US-08-416-711-4
US-09-356-497-4
US-09-134-001C-4654
US-09-198-452A-797
US-09-232-225-21
US-09-292-225-15
US-09-292-225-18
US-09-107-532A-6086
US-09-543-681A-4524
US-09-266-965-101
US-08-469-260A-84
US-08-488-446-84
US-08-467-344A-84
US-09-107-532A-5965
US-09-489-039A-9709
US-08-983-607-23
US-09-489-039A-9477
US-08-591-468-7

ALIGNMENTS

RESULT 1
US-08-416-711-3
; Sequence 3, Application US/08416711
; Patent No. 6017538
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: BOUHAROUN-TAYOUN, HASNAQ
; APPLICANT: OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
; TITLE OF INVENTION: PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-416-711-3


```

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 100.0%; Score 150; DB 4; Length 64;
Best Local Similarity 100.0%; Pred. No. 7.4e-16;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKESYDYILGWEEFGGVPPEHKKEEN 27
DB 18 AKESYDYILGWEEFGGVPPEHKKEEN 44

RESULT 5
US-09-252-991A-32073
; Sequence 32073, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32073
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32073

Query Match 33.7%; Score 50.5; DB 4; Length 597;
Best Local Similarity 44.0%; Pred. No. 15;
Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 2 KEASSYDYILGWEEFGGVPPEHKKE 25
DB 563 KPFNAVKHSIGWGDWGVDPDKLE 587

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 64 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-356-497-1

Query Match 100.0%; Score 150; DB 4; Length 64;
Best Local Similarity 100.0%; Pred. No. 7.4e-16;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKESYDYILGWEEFGGVPPEHKKEEN 27
DB 18 AKESYDYILGWEEFGGVPPEHKKEEN 44

RESULT 5
US-09-252-991A-32073
; Sequence 32073, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32073
; LENGTH: 597
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32073

Query Match 33.7%; Score 50.5; DB 4; Length 597;
Best Local Similarity 44.0%; Pred. No. 15;
Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 2 KEASSYDYILGWEEFGGVPPEHKKE 25
DB 563 KPFNAVKHSIGWGDWGVDPDKLE 587
```

```

RESULT 6
US-09-292-225-41
; Sequence 41, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 41
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-41

Query Match 32.7%; Score 49; DB 4; Length 490;
Best Local Similarity 43.5%; Pred. No. 20;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

QY 6 SYDYILGWE--FCGGVPPEHKKE 26
DB 206 TYDHGGWENFGHNAPLYKRPD 228
```

```

RESULT 7
US-09-292-225-35
; Sequence 35, Application US/09292225
; Patent No. 6455686
; GENERAL INFORMATION:
; APPLICANT: McCall, Catherine A.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
; FILE REFERENCE: AL-2-C3
; CURRENT APPLICATION NUMBER: US/09/292,225
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: 60/098,909
; EARLIER FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/085,295
; EARLIER FILING DATE: 1998-05-13
; EARLIER APPLICATION NUMBER: 60/098,565
; EARLIER FILING DATE: 1998-04-17
; EARLIER APPLICATION NUMBER: 09/062,013
; EARLIER FILING DATE: 1998-04-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Dermatophagoides farinae
US-09-292-225-35

Query Match 32.7%; Score 49; DB 4; Length 509;
Best Local Similarity 43.5%; Pred. No. 21;
Matches 10; Conservative 4; Mismatches 7; Indels 2; Gaps 1;
```

QY 6 SYDYILGWE--FGGVPPEHKKEE 26
 Db 225 TYDHGGWENVFHGNAPLYKRPD 247

RESULT 8
 US-09-292-225-38
 ; Sequence 38, Application US/09292225
 ; Patent No. 6455686
 ; GENERAL INFORMATION:
 ; APPLICANT: McCall, Catherine A.
 ; APPLICANT: Hunter, Shirley Wu
 ; APPLICANT: Weber, Eric R.
 ; TITLE OF INVENTION: NOVEL DERMATOPHAGOIDES NUCLEIC ACID MOLECULES, PROTEINS
 ; TITLE OF INVENTION: AND USES THEREOF
 ; FILE REFERENCE: AL-2-C3
 ; CURRENT APPLICATION NUMBER: US/09/292,225
 ; CURRENT FILING DATE: 1999-04-15
 ; EARLIER APPLICATION NUMBER: 60/098,909
 ; EARLIER FILING DATE: 1998-09-02
 ; EARLIER APPLICATION NUMBER: 60/085,295
 ; EARLIER FILING DATE: 1998-05-13
 ; EARLIER APPLICATION NUMBER: 60/098,565
 ; EARLIER FILING DATE: 1998-04-17
 ; EARLIER APPLICATION NUMBER: 09/062,013
 ; EARLIER FILING DATE: 1998-04-17
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 38
 ; TYPE: PRT
 ; LENGTH: 509
 ; ORGANISM: Dermatophagoides farinae
 US-09-292-225-38

Query Match 32.7%; Score 49; DB 4; Length 509;
 Best Local Similarity 43.5%; Pred. No. 21;
 Matches 10; Conservative 4; Mismatches 7; Indels 7; Gaps 1;

QY 6 SYDYILGWE--FGGVPPEHKKEE 26
 Db 225 TYDHGGWENVFHGNAPLYKRPD 247

RESULT 9
 US-09-386-642-53
 ; Sequence 53, Application US/09386642
 ; Patent No. 6420157
 ; GENERAL INFORMATION:
 ; APPLICANT: Darrow, Andrew
 ; APPLICANT: Qi, Jensen
 ; APPLICANT: Andrade-Gordon, Patricia
 ; TITLE OF INVENTION: Zymogen Activation System
 ; FILE REFERENCE: ORT-1028
 ; CURRENT APPLICATION NUMBER: US/09/386,642
 ; CURRENT FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 53
 ; TYPE: PRT
 ; LENGTH: 306
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene of
 ; OTHER INFORMATION: human protease F in CFEK2 zymogen vector
 US-09-386-642-53

Query Match 32.0%; Score 48; DB 4; Length 306;
 Best Local Similarity 44.4%; Pred. No. 17;
 Matches 8; Conservative 6; Mismatches 4; Indels 4; Gaps 0;

QY 10 ILGWFEFGGVPPEHKKEEN 27
 Db 12 LLGTTFGCGVPDYKDDDD 29

RESULT 10
 US-09-386-642-12
 ; Sequence 12, Application US/09386642
 ; Patent No. 6420157
 ; GENERAL INFORMATION:
 ; APPLICANT: Darrow, Andrew
 ; APPLICANT: Qi, Jensen
 ; APPLICANT: Andrade-Gordon, Patricia
 ; TITLE OF INVENTION: Zymogen Activation System
 ; FILE REFERENCE: ORT-1028
 ; CURRENT APPLICATION NUMBER: US/09/386,642
 ; CURRENT FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 60
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12
 ; TYPE: PRT
 ; LENGTH: 319
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
 ; OTHER INFORMATION: with homo sapien serine protease catalytic domain
 US-09-386-642-12

Query Match 32.0%; Score 48; DB 4; Length 319;
 Best Local Similarity 44.4%; Pred. No. 18;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 10 ILGWFEFGGVPPEHKKEEN 27
 Db 12 LLGTTFGCGVPDYKDDDD 29

RESULT 11
 US-09-252-991A-22343
 ; Sequence 22343, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 22343
 ; LENGTH: 560
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-22343

Query Match 32.0%; Score 48; DB 4; Length 560;
 Best Local Similarity 56.2%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 4 ASSDYILGWEFGGV 19
 Db 379 AVSNTYTLNWDGSGV 394

RESULT 12
 US-09-252-991A-19016
 ; Sequence 19016, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 19016
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-19016

Query Match 32.0%; Score 48; DB 4; Length 394;
 Best Local Similarity 56.2%; Pred. No. 34;
 Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19016
LENGTH: 671
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19016

Query Match 32.0%; Score 48; DB 4; Length 671;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 10 ILGWFGGV 19
212 LLGWSFGGL 221

RESULT 13
US-09-134-001C-3592
Sequence 3592, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134.001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3592
LENGTH: 593
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3592

Query Match 31.7%; Score 47.5; DB 4; Length 593;
Best Local Similarity 52.9%; Pred. No. 43;
Matches 9; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy 3 EASSYDYL-GWFFGG 18
478 QANAYDIVLNGYELGGG 494

RESULT 14
US-09-543-681A-6335
Sequence 6335, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 6335
LENGTH: 168
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-6335

Query Match 31.3%; Score 47; DB 4; Length 168;
Best Local Similarity 45.8%; Pred. No. 13;
Matches 11; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

Qy 1 AKEASSYDYLGWFFGGVPEHKK 24
Db 82 AKDSFYIYI--WNYGDVLEGGK 103

RESULT 15
US-09-100-664A-9
Sequence 9, Application US/09100664A
Patent No. 6057129
GENERAL INFORMATION:
APPLICANT: YOUNG, MICHAEL W.
APPLICANT: KLOSS, BRIAN
APPLICANT: BLAU, JUSTIN
APPLICANT: PRICE, JEFFREY
TITLE OF INVENTION: A NOVEL CLOCK GENE AND METHODS OF USE
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,664A
FILING DATE: 19-JUN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-221
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-100-664A-9

Query Match 31.3%; Score 47; DB 3; Length 416;
Best Local Similarity 29.7%; Pred. No. 35;
Matches 11; Conservative 6; Mismatches 8; Indels 12; Gaps 2;

Qy 2 KEASSYDYLGW---BFGGG-----VPEHKKEE 26
Db 279 RQGSFYDYVDNNMLKFGAARNPEDVDRREREERE 315

Search completed: September 22, 2004, 18:11:01
Job time : 13.719 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:09:13 ; Search time 58.2397 Seconds
(without alignments)
148.878 Million cell updates/sec

Title: US-10-774-602-12
Perfect score: 150
Sequence: 1 AKEASSYDYLGWFGGVPPEHKEEN 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	150	100.0	27 14	US-10-294-770-3	Sequence 3, Appl
2	150	100.0	27 14	US-10-294-770-12	Sequence 12, Appl
3	150	100.0	27 16	US-10-238-741-3	Sequence 3, Appl
4	150	100.0	27 16	US-10-774-602-3	Sequence 3, Appl
5	150	100.0	27 16	US-10-774-602-12	Sequence 12, Appl
6	150	100.0	64 14	US-10-294-770-1	Sequence 1, Appl
7	150	100.0	64 16	US-10-238-741-1	Sequence 1, Appl
8	150	100.0	64 16	US-10-774-602-1	Sequence 1, Appl
9	58	38.7	596 12	US-10-282-122A-51159	Sequence 51159, A
10	58	38.7	599 12	US-10-282-122A-49798	Sequence 49798, A
11	58	38.7	600 12	US-10-282-122A-48020	Sequence 48020, A
12	58	38.7	600 12	US-10-282-122A-50005	Sequence 50005, A
13	57	38.0	382 12	US-10-282-122A-43296	Sequence 43296, A
14	57	38.0	382 14	US-10-210-115-33	Sequence 33, Appl
15	57	38.0	382 15	US-10-369-493-691	Sequence 691, App

16	55	36.7	247	12	US-10-335-977-4888	Sequence 4888, Ap
17	55	36.7	479	9	US-09-881-752A-14	Sequence 14, Appl
18	55	36.7	479	12	US-10-335-977-4890	Sequence 4890, Ap
19	55	36.7	486	12	US-10-335-977-4891	Sequence 4891, Ap
20	54	36.0	432	16	US-10-437-963-123350	Sequence 123350, A
21	52.5	35.0	86	16	US-10-767-701-51357	Sequence 51357, A
22	52.5	35.0	111	14	US-10-091-300-35	Sequence 35, Appl
23	52.5	35.0	111	14	US-10-091-300-51	Sequence 51, Appl
24	52.5	35.0	594	12	US-10-282-122A-65399	Sequence 65399, A
25	52.5	35.0	602	12	US-10-282-122A-66049	Sequence 66049, A
26	52	34.7	378	15	US-10-369-493-360	Sequence 360, App
27	52	34.7	858	15	US-10-369-493-5536	Sequence 5536, Ap
28	52	34.7	858	15	US-10-369-493-5537	Sequence 5537, Ap
29	51.5	34.3	269	9	US-09-888-721-36	Sequence 36, Appl
30	51.5	34.3	282	9	US-09-888-721-38	Sequence 38, Appl
31	51.5	34.3	287	9	US-09-888-721-40	Sequence 40, Appl
32	51.5	34.3	251	9	US-09-888-721-44	Sequence 44, Appl
33	51.5	34.3	296	9	US-09-888-721-42	Sequence 42, Appl
34	51	34.0	66	16	US-10-437-963-106497	Sequence 106497, A
35	51	34.0	263	14	US-10-156-761-9233	Sequence 9233, Ap
36	51	34.0	404	16	US-10-437-963-103597	Sequence 103597, A
37	51	34.0	605	16	US-10-437-963-139769	Sequence 139769, A
38	50.5	33.7	588	9	US-09-815-242-12553	Sequence 12553, A
39	50.5	33.7	588	9	US-09-815-242-12816	Sequence 12816, A
40	50.5	33.7	588	12	US-10-282-122A-44049	Sequence 44049, A
41	50.5	33.7	590	9	US-09-815-242-5344	Sequence 5344, Ap
42	50	33.3	59	12	US-10-424-599-254448	Sequence 254448, A
43	50	33.3	174	12	US-10-335-977-7134	Sequence 7134, Ap
44	50	33.3	238	10	US-09-880-748-2067	Sequence 2067, Ap
45	50	33.3	238	12	US-10-293-418-2067	Sequence 2067, Ap

ALIGNMENTS

RESULT 1
US-10-294-770-3
; Sequence 3, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR FILING DATE: 2002-11-15
; PRIOR FILING DATE: 1999-07-19
; PRIOR FILING DATE: 1995-08-08
; PRIOR FILING DATE: 1995-08-08
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-294-770-3

Query Match 100.0%; Score 150; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLGWFGGVPPEHKEEN 27
Db 1 AKEASSYDYLGWFGGVPPEHKEEN 27

RESULT 2
US-10-294-770-12
; Sequence 12, Application US/10294770
; Publication No. US20030161840A1
; GENERAL INFORMATION:

; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 230759USOCIP
; CURRENT APPLICATION NUMBER: US/10/294,770
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; -294-770-12

Query Match 100.0%; Score 150; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 3
US-10-238-741-3
; Sequence 3, Application US/10238741
; Publication No. US20040096466A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; BOUHAROUN-TAYOUN, HASNAQ
; OEUVRAY, CLAUDE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCELLELAND, MAIER & NEUSTADT, P.C.
; STREET: 1755 S. JERFPERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/238,741
; FILING DATE: 09-Nov-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,497
; FILING DATE: 19-Jul-1999
; APPLICATION NUMBER: US/08/416,711
; FILING DATE: 08-AUG-1995
; APPLICATION NUMBER: PCT/FR93/01024
; FILING DATE: 18-OCT-1993
; APPLICATION NUMBER: FR 92/12488
; FILING DATE: 19-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-238-741-3

Query Match 100.0%; Score 150; DB 16; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 4
US-10-774-602-3
; Sequence 3, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-774-602-3

Query Match 100.0%; Score 150; DB 16; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPHEHKEEN 27
DB 1 AKEASSYDYLGWFGGVPHEHKEEN 27

RESULT 5
US-10-774-602-12
; Sequence 12, Application US/10774602
; Publication No. US20040141987A1
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
; FILE REFERENCE: 248791USODIV
; CURRENT APPLICATION NUMBER: US/10/774,602
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: US 09/356,947
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: US 10/238,741
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: US 08/416,711
; PRIOR FILING DATE: 1995-08-08
; PRIOR APPLICATION NUMBER: PCT/FR93/01024
; PRIOR FILING DATE: 1993-10-18

; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 12
 ; LENGTH: 27
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Peptide
 US-10-774-602-12

 Query Match 100.0%; Score 150; DB 16; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.9e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 AKEASSYDYLWFGGVPHEKKEN 27
 |||||
 DB 1 AKEASSYDYLWFGGVPHEKKEN 27
 |||||

 RESULT 6
 US-10-294-770-1
 ; Sequence 1, Application US/10294770
 ; Publication No. US20030161840A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUILHE, PIERRE
 ; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 ; FILE REFERENCE: 230759USOCIP
 ; CURRENT APPLICATION NUMBER: US/10/294,770
 ; CURRENT FILING DATE: 2002-11-15
 ; PRIOR APPLICATION NUMBER: US 09/356,947
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR FILING DATE: 1995-08-08
 ; PRIOR APPLICATION NUMBER: PCT/FR93/01024
 ; PRIOR FILING DATE: 1993-10-18
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 64
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 US-10-294-770-1

 Query Match 100.0%; Score 150; DB 14; Length 64;
 Best Local Similarity 100.0%; Pred. No. 4.8e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 AKEASSYDYLWFGGVPHEKKEN 27
 |||||
 DB 18 AKEASSYDYLWFGGVPHEKKEN 44
 |||||

 RESULT 7
 US-10-238-741-1
 ; Sequence 1, Application US/10238741
 ; Publication No. US20040096466A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUILHE, PIERRE
 ; BOUHAROUN-TAYOUN, HASNAQ
 ; OEUVRAY, CLAUDE
 ; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING
 ; PROTECTIVE ANTIBODIES
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ; P.C.
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
 ; CITY: ARLINGTON
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/238,741
 ; FILING DATE: 09-Nov-2002
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/356,497
 ; FILING DATE: 19-Jul-1999
 ; APPLICATION NUMBER: US/08/416,711
 ; FILING DATE: 08-AUG-1995
 ; APPLICATION NUMBER: PCT/FR93/01024
 ; FILING DATE: 18-OCT-1993
 ; APPLICATION NUMBER: FR 92/12488
 ; FILING DATE: 19-OCT-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F.
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 660-085-0 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-413-3000
 ; TELEFAX: 703-413-2220
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 64 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-238-741-1

 Query Match 100.0%; Score 150; DB 16; Length 64;
 Best Local Similarity 100.0%; Pred. No. 4.8e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 AKEASSYDYLWFGGVPHEKKEN 27
 |||||
 DB 18 AKEASSYDYLWFGGVPHEKKEN 44
 |||||

 RESULT 8
 US-10-774-602-1
 ; Sequence 1, Application US/10774602
 ; Publication No. US20040141987A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DRUILHE, PIERRE
 ; TITLE OF INVENTION: PLASMODIUM FALCIPARUM ANTIGENS INDUCING PROTECTIVE ANTIBODIES
 ; FILE REFERENCE: 248791USODIV
 ; CURRENT APPLICATION NUMBER: US/10/774,602
 ; CURRENT FILING DATE: 2004-02-10
 ; PRIOR APPLICATION NUMBER: US 09/356,947
 ; PRIOR FILING DATE: 1999-07-19
 ; PRIOR APPLICATION NUMBER: US 10/238,741
 ; PRIOR FILING DATE: 2002-09-11
 ; PRIOR APPLICATION NUMBER: US 08/416,711
 ; PRIOR FILING DATE: 1995-08-08
 ; PRIOR APPLICATION NUMBER: PCT/FR93/01024
 ; PRIOR FILING DATE: 1993-10-18
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patent in version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 64
 ; TYPE: PRT
 ; ORGANISM: Plasmodium falciparum
 US-10-774-602-1

 Query Match 100.0%; Score 150; DB 16; Length 64;
 Best Local Similarity 100.0%; Pred. No. 4.8e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 AKEASSYDYLWFGGVPHEKKEN 27
 |||||
 DB 18 AKEASSYDYLWFGGVPHEKKEN 44
 |||||


```

RESULT 12
US/10-282-122A-50005
// Sequence 50005, Application US/10282122A
// Publication No. US20040029129A1
// GENERAL INFORMATION:
// APPLICANT: Wang, Liangsu
// APPLICANT: Zamudio, Carlos
// APPLICANT: Malone, Cheryl
// APPLICANT: Haselbeck, Robert
// APPLICANT: Ohlsen, Kari
// APPLICANT: Zyskind, Judith
// APPLICANT: Wall, Daniel
// APPLICANT: Trawick, John
// APPLICANT: Carr, Grant
// APPLICANT: Yamamoto, Robert
// APPLICANT: Forsyth, R.
// APPLICANT: Xu, H.
// TITLE OF INVENTION: Identification of E
// FILE REFERENCE: ELIPIRA.034A
// CURRENT APPLICATION NUMBER: US/10/282,1
// CURRENT FILING DATE: 2003-02-20
// PRIOR APPLICATION NUMBER: 60/191,078
// PRIOR FILING DATE: 2000-03-21
// PRIOR APPLICATION NUMBER: 60/206,848
// PRIOR FILING DATE: 2000-05-23
// PRIOR APPLICATION NUMBER: 60/207,727
// PRIOR FILING DATE: 2000-05-26
// PRIOR APPLICATION NUMBER: 60/230,335
// PRIOR FILING DATE: 2000-09-06
// PRIOR APPLICATION NUMBER: 60/230,347
// PRIOR FILING DATE: 2000-09-09
// PRIOR APPLICATION NUMBER: 60/242,578
// PRIOR FILING DATE: 2000-10-23
// PRIOR APPLICATION NUMBER: 60/253,625
// PRIOR FILING DATE: 2000-11-27
// PRIOR APPLICATION NUMBER: 60/257,931
// PRIOR FILING DATE: 2000-12-22
// PRIOR APPLICATION NUMBER: 60/267,636

```

RESULT 12
US-10-282-122A-50005
Sequence 50005, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELTRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636

Query Match 38.0%; Score 57; DB 12; Length 382;
Best Local Similarity 46.7%; Pred. No. 8.6;

Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;
 QY 1 AKE---ASSVDYILG-WFEGGGVPEHKKEE 26
 Db 159 AKEVTTAEAYSWTQGSWTLTGLPEAKKD 188
 Search completed: September 22, 2004, 18:34:04
 Job time : 59.2397 secs

RESULT 14
 US-10-210-115-33
 ; Sequence 33, Application US/10210115
 ; Publication No. US20030129708A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ptitsyn, Leonid R
 ; APPLICANT: Smirnov, Sergey V
 ; APPLICANT: Altman, Irina B
 ; APPLICANT: No. US20030129708A1ikova, Anna E
 ; APPLICANT: Kotliarova, Veronika A
 ; APPLICANT: Gusyatin, Mikhail M
 ; APPLICANT: Rostova, Yulia G
 ; APPLICANT: Yampolskaya, Tatiana A
 ; TITLE OF INVENTION: New Mutant Carbamoylphosphate Synthetase and Method for Producing
 ; TITLE OF INVENTION: Compounds Derived from Carbamoylphosphate
 ; FILE REFERENCE: 226350US0
 ; CURRENT APPLICATION NUMBER: US/10/210,115
 ; CURRENT FILING DATE: 2002-11-12
 ; PRIOR APPLICATION NUMBER: RU 2001-121697
 ; PRIOR FILING DATE: 2001-08-03
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent version 3.1
 ; SEQ ID NO 33
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-210-115-33

Query Match 38.0%; Score 57; DB 14; Length 382;
 Best Local Similarity 46.7%; Pred. No. 8.6;
 Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;
 QY 1 AKE---ASSVDYILG-WFEGGGVPEHKKEE 26
 Db 159 AKEVTTAEAYSWTQGSWTLTGLPEAKKD 188

RESULT 15
 US-10-369-493-691
 ; Sequence 691, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 691
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-369-493-691

Query Match 38.0%; Score 57; DB 15; Length 382;
 Best Local Similarity 46.7%; Pred. No. 8.6;
 Matches 14; Conservative 4; Mismatches 8; Indels 4; Gaps 2;
 QY 1 AKE---ASSVDYILG-WFEGGGVPEHKKEE 26

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 43.3719 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-14
Perfect score: 204
Sequence: 1 MLSHLYVSKDKENISKEND.....VLDEKEEABETEELK 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	329	5 Q9NFV9	Q9NFV9 plasmodium
2	204	100.0	354	5 Q25995	Q25995 plasmodium
3	204	100.0	354	5 Q81J55	Q81J55 plasmodium
4	204	100.0	361	5 Q95PI5	Q95PI5 plasmodium
5	204	100.0	379	5 Q25705	Q25705 plasmodium
6	204	100.0	379	5 Q25706	Q25706 plasmodium
7	204	100.0	379	5 Q9UK64	Q9UK64 plasmodium
8	204	100.0	380	5 Q26019	Q26019 plasmodium
9	194	95.1	346	5 Q9U0G0	Q9U0G0 plasmodium
10	97	47.5	64	5 Q15762	Q15762 plasmodium
11	94.5	46.3	424	5 Q81J53	Q81J53 plasmodium
12	89	43.6	697	5 Q81J52	Q81J52 plasmodium
13	87	42.6	343	5 Q8SZT9	Q8SZT9 drosophila
14	87	42.6	1147	5 Q9VYT3	Q9VYT3 drosophila
15	85.5	41.9	406	4 Q8TC47	Q8TC47 homo sapien
16	85.5	41.9	406	4 Q8TC90	Q8TC90 homo sapien

17	84	41.2	405	5	Q81J48	Q81J48 plasmodium
18	82	40.2	1015	5	Q86AG0	Q86AG0 dictyosteli
19	82	40.2	3134	5	Q25994	Q25994 plasmodium
20	82	40.2	3144	5	Q9GTK4	Q9GTK4 plasmodium
21	82	40.2	5890	5	Q81K84	Q81K84 plasmodium
22	81.5	40.0	142	11	Q8C7N3	Q8C7N3 mus musculu
23	80	39.2	232	10	Q9ST20	Q9ST20 brassica ca
24	80	39.2	662	5	Q81AW6	Q81AW6 plasmodium
25	80	39.2	806	5	Q8T2A3	Q8T2A3 dictyosteli
26	80	39.2	874	5	Q81CL6	Q81CL6 plasmodium
27	79.5	39.0	152	10	Q65745	Q65745 cicer ariet
28	79	38.7	204	5	Q81518	Q81518 plasmodium
29	79	38.7	477	13	Q8AYN2	Q8AYN2 gallus gall
30	79	38.7	715	10	Q9LD59	Q9LD59 arabidopsis
31	79	38.7	741	16	Q9JTK3	Q9JTK3 neisseria m
32	79	38.7	1192	4	Q9UL68	Q9UL68 homo sapien
33	79	38.7	1989	5	Q814U7	Q814U7 plasmodium
34	79	38.7	2924	5	Q25733	Q25733 plasmodium
35	78.5	38.5	395	4	Q8NDU3	Q8NDU3 homo sapien
36	78.5	38.5	695	11	Q8CDJ9	Q8CDJ9 mus musculu
37	78.5	38.5	1331	11	Q9SPQ2	Q9SPQ2 mus musculu
38	78	38.2	331	10	Q8LNT8	Q8LNT8 oryza sativ
39	78	38.2	400	11	Q7TP09	Q7TP09 rattus norv
40	78	38.2	755	11	Q8OW59	Q8OW59 rattus norv
41	78	38.2	772	11	Q8BTT6	Q8BTT6 mus musculu
42	78	38.2	910	12	Q9Q9I5	Q9Q9I5 avian adeno
43	78	38.2	1781	4	Q9UKX0	Q9UKX0 homo sapien
44	78	38.2	1890	4	Q9UKW3	Q9UKW3 homo sapien
45	78	38.2	2072	4	Q8WYB5	Q8WYB5 homo sapien

ALIGNMENTS

RESULT 1

ID Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (fragment).
GN MSP3.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252287; CAB85901.1; -;
FT NON_TER 1 1
FT NON_TER 329 329
SQ SEQUENCE 329 AA; C5B045DB5E21A159 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 5e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLSHLYVSKDKENISKENDVLDKEEABETEELK 41
Db 201 MLSHLYVSKDKENISKENDVLDKEEABETEELK 241

RESULT 2

Q25995
ID Q25995 PRELIMINARY; PRT; 354 AA.
AC Q25995;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)

```

DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DE Antigen.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=NF54;
RX MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
RA Thompson J.K., Marshall V.M., Coppal R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
RT Plasmodium falciparum merozoites";
RL Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L28825; AAC09377.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.3e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
|||||
Db 211 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 251

RESULT 3
Q8IJ55 PRELIMINARY; PRT; 354 AA.
AC Q8IJ55;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DE Merozoite surface protein 3.
GN PF10_0345.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J.J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511(2002).
DR EMBL; AB014834; AAN35542.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.3e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
|||||
Db 211 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 251

RESULT 4
Q95PI5 PRELIMINARY; PRT; 361 AA.
AC Q95PI5;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=FVO;
RC Hisaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
RA "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus
RA Monkeys.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY044180; AAK94780.1; -.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 361;
Best Local Similarity 100.0%; Pred. No. 5.4e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
|||||
Db 217 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 257

RESULT 5
Q25705 PRELIMINARY; PRT; 379 AA.
AC Q25705;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=X1;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08851; AAC47831.1; -.
SQ SEQUENCE 379 AA; 43344 MW; DC7AF106887C9AA0 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 41
|||||
Db 237 MLSHLYVSSKDKENISKENDVLDKEKEAEETEEELK 277

RESULT 6
Q25706 PRELIMINARY; PRT; 379 AA.
AC Q25706;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Polymorphic antigen.
```

```

OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FC27;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08852; AAC47832.1; -.
SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 41
237 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 277

RESULT 7
Q9U6C4 PRELIMINARY; PRT; 379 AA.
AC Q9U6C4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
GN MSP-3.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RA Li X.R., Yu X.B., Luo S.H., Shan Z.X., Fang J.M., Xu J.;
RT "Sequence of Plasmodium falciparum secreted polymorphic antigen
RT gene.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188190; AAF04099.1; -.
SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 41
237 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 277

RESULT 8
Q26019 PRELIMINARY; PRT; 380 AA.
AC Q26019;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
RT Plasmodium falciparum merozoites.";

```

```

RL Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=FC27;
RC MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L07944; AAC09378.1; -.
DR PDB; 1PSM; 07-FEB-95.
KW SIGNAL.
FT SIGNAL.
FT CHAIN 1 25 POTENTIAL.
FT CHAIN 26 380 POLYMORPHIC ANTIGEN.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 204; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.7e-12;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 41
238 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 278

RESULT 9
Q9U0G0 PRELIMINARY; PRT; 346 AA.
AC Q9U0G0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Pragmat).
GN MSP3.
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252286; CAB65754.1; -.
FT NON_TER 1 1
FT NON_TER 346 346
SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;

Query Match 95.1%; Score 194; DB 5; Length 346;
Best Local Similarity 95.1%; Pred. No. 4.7e-11;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 41
214 MLSHLYVSSKDKENISKENDVLDKEEEAEETEEELK 254

RESULT 10
O15762 PRELIMINARY; PRT; 64 AA.
AC O15762;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP-3.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;

```

XX	MEDLINE=94348085; PubMed=8068948;
RA	Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA	Aikawa M., Filgueira M.C., Tartar A., Druilhé P.;
RT	"Merozoite surface protein-3: a malaria protein inducing antibodies
RT	that promote Plasmodium falciparum killing by cooperation with blood
RT	monocytes.";
RL	Blood 84:11594-1602(1994).
RN	[2]
RC	SEQUENCE FROM N.A.
RP	STRAIN=NF54;
RA	Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA	Aikawa M., Filgueira M.-C., Tartar A., Druilhé P.;
RL	Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF024624; AAB81610.1; -.
FT	NON_TER 1 1
FT	NON_TER 64 64
SQ	SEQUENCE 64 AA; 7326 MW; 02E0ABF614C15A3B CRC64;
Query Match	47.5%; Score 97; DB 5; Length 64;
Best Local Similarity	95.0%; Pred. No. 0.02;
Matches 19; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
QY	1 MLSHLYVSSKDKENISKEND 20
DG	
DB	45 MLSHLYVSSKDKENISKENE 64
RESULT 11	
ID Q8IJ53	PRELIMINARY; PRT; 424 AA.
AC Q8IJ53;	
DT 01-MAR-2003 (TrEMBLrel. 23, Created)	
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)	
DE Hypothetical protein.	
DE PF10_0347.	
OS Plasmodium falciparum (isolate 3D7).	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.	
OX NCBI_TaxID=36329;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=3D7;	
RX MEDLINE=22255705; PubMed=12368864;	
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,	
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,	
RA Chan J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,	
RA Chen M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguoli S.,	
RA Portea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,	
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,	
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,	
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,	
RA Fraser C.M., Barrell B.;	
RT "Genome sequence of the human malaria parasite Plasmodium	
RT falciparum."	
RL Nature 419:498-511(2002).	
DR EMBL; AE014834; AAN35544.1; -.	
KW Hypothetical protein.	
SQ SEQUENCE 424 AA; 48196 MW; 4D2BD2C9A9F3B619 CRC64;	
Query Match	46.3%; Score 94.5; DB 5; Length 424;
Best Local Similarity	51.2%; Pred. No. 0.19;
Matches 21; Conservative	10; Mismatches 9; Indels 1; Gaps 1;
QY	1 MLSHLYVSSKDKENISKENDVDLDEKEEAETEEELEEK 41
DG	: : : : : : :
DB	285 VLRSVNLTWDKDIKVENEDVKDEKDEDDDEE-BEEKYENE 324
RESULT 12	
ID Q8IJ52	PRELIMINARY; PRT; 697 AA.
AC Q8IJ52;	
DT 01-MAR-2003 (TrEMBLrel. 23, Created)	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 19:00:42 ; Search time 39.876 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-11

Perfect score: 124

Sequence: 1 YEKAKYAKQANQAVLKAEASSYD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Optimal number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	93.5	64	2 AAR54684	Aar54684 Plasmodium
2	116	93.5	64	6 ABR41813	AbR41813 Plasmodium
3	115	92.7	28	6 ABR41814	AbR41814 Plasmodium
4	106	85.5	31	6 ABR41822	AbR41822 Plasmodium
5	103	83.1	31	6 ABR41823	AbR41823 Plasmodium
6	99	79.8	31	6 ABR41824	AbR41824 Plasmodium
7	97	78.2	31	6 ABR41825	AbR41825 Plasmodium
8	96	77.4	23	2 AAR54685	Aar54685 Plasmodium
9	91	73.4	31	6 ABR41826	AbR41826 Plasmodium
10	88	71.0	31	6 ABR41827	AbR41827 Plasmodium
11	84	67.7	31	6 ABR41828	AbR41828 Plasmodium
12	80	64.5	31	6 ABR41829	AbR41829 Plasmodium
13	69	55.6	31	6 ABR41830	AbR41830 Plasmodium
14	67	54.0	31	6 ABR41833	AbR41833 Plasmodium
15	65	52.4	31	6 ABR41832	AbR41832 Plasmodium
16	63	50.8	31	6 ABR41831	AbR41831 Plasmodium
17	61	49.2	31	6 ABR41834	AbR41834 Plasmodium
18	56	45.2	31	6 ABR41835	AbR41835 Plasmodium
19	55	44.4	752	5 ABB54485	Abb54485 Plasmodium
20	53	42.7	463	4 ABB59725	Abb59725 Plasmodium
21	52	41.9	31	6 ABR41836	AbR41836 Plasmodium
22	51	41.1	347	3 AAG23630	Aag23630 Plasmodium
23	51	41.1	368	3 AAG23629	Aag23629 Plasmodium
24	51	41.1	420	3 AAG23628	Aag23628 Plasmodium
25	50	40.3	28	2 AAR90177	Aar90177 Plasmodium

ALIGNMENTS

RESULT 1

AAR54684
ID AAR54684 standard; protein; 64 AA.

XX AAR54684;

XX 25-MAR-2003 (revised)

DT 15-DEC-1994 (first entry)

DT 15-DEC-1994 (first entry)

DE Plasmodium falciparum merozoite 48kD surface protein fragment I.

XX Merozoite 48kD surface antigen; immunogen; vaccine; malaria;

XX cytophilic antibody.

XX Plasmodium falciparum.

XX Plasmodium falciparum.

PN WO9409140-A1.

XX 28-APR-1994.

XX 18-OCT-1993; 93WO-FR001024.

XX 19-OCT-1992; 92FR-00012488.

PR (INSP) INST PASTEUR.

PI Drulilhe P, Bouharoun-Tayoun H, Oeuvray C;

XX WPI; 1994-151325/18.

DR N-PSDB; AAQ64677.

XX New Plasmodium antigen generating cytophilic - mimicking protection induced by long term exposure to the parasite, useful in vaccination against and diagnosis of malaria, and prodn of antibodies for passive immunisation.

PS Claim 3; Page 36; 52pp; French.

CC This sequence is a preferred fragment of the 48kD merozoite surface antigen of P.falciparum which is useful as an immunogen to generate antibodies (Abs). The Abs can induce a cytotoxic reaction against the erythrocyte stage of the parasite. Specific subfragments of AAR54684 are separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 64 AA;

SQ Query Match 93.5%; Score 116; DB 2; Length 64;

```

Best Local Similarity 92.0%; Pred. No. 6.4e-10;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
   :|:|||||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 2
ABR41813
ID ABR41813 standard; protein; 64 AA.
XX
AC ABR41813;
XX
DT 11-AUG-2003 (first entry)
XX
DE Merozoite surface antigen MSP-3 peptide (aa194-257).
XX
KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
immunotherapy; vaccine.
XX
OS Plasmodium falciparum.
XX
FN GB2378949-A.
XX
PD 26-FEB-2003.
XX
PF 16-AUG-2001; 2001GB-00020057.
XX
PR 16-AUG-2001; 2001GB-00020057.
XX
(DZIE/) DZIEGIEL M S H.
PA (LUND/) LUNDQUIST R.
PA (NIEL/) NIELSEN L K.
XX
PI Dziegiel MSH, Lundquist R, Nielsen LK;
XX
DR WPI; 2003-303108/30.
XX
Novel recombinant human antibody specific for merozoite surface protein-3
antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
a malarial disease.
XX
PS Disclosure; Fig 3; 75pp; English.
XX
The present sequence is that of amino acids 194-257 of the merozoite
surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
invention relates to the construction of recombinant human antibodies
against MSP-3, in particular MSP-3(194-257) antigen. The antibodies (see
also ABR41816-21) are able to passively induce naturally occurring non-
sterile malaria immunity (termed premunition) and exert their effect via
the antibody-dependent cellular inhibition of growth (ADCI) mechanism.
They can be used in the diagnosis and treatment of malaria. The invention
also provides a method of screening a phage library with target
molecules, such as MSP-3 antigen, bound to beads, preferably magnetic
microspheres
XX
SQ Sequence 64 AA;

Query Match 93.5%; Score 116; DB 6; Length 64;
Best Local Similarity 92.0%; Pred. No. 6.4e-10;
Matches 23; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
   :|:|||||
Db 1 HERAKNAYOKANQAVLKAKEASSYD 25

RESULT 3
ABR41822
ID ABR41822 standard; peptide; 31 AA.
XX
AC ABR41822;
XX
DT 11-AUG-2003 (first entry)
XX
DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
immunotherapy; vaccine; epitope.
XX
OS Plasmodium falciparum.

```

```

ABR41814
ID ABR41814 standard; protein; 28 AA.
XX
AC ABR41814;
XX
DT 11-AUG-2003 (first entry)
XX
DE Merozoite surface antigen MSP-3 peptide (aa190-217).
XX
KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
immunotherapy; vaccine.
XX
OS Plasmodium falciparum.
XX
FN GB2378949-A.
XX
PD 26-FEB-2003.
XX
PF 16-AUG-2001; 2001GB-00020057.
XX
PR 16-AUG-2001; 2001GB-00020057.
XX
(DZIE/) DZIEGIEL M S H.
PA (LUND/) LUNDQUIST R.
PA (NIEL/) NIELSEN L K.
XX
PI Dziegiel MSH, Lundquist R, Nielsen LK;
XX
DR WPI; 2003-303108/30.
XX
Novel recombinant human antibody specific for merozoite surface protein-3
antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
a malarial disease.
XX
PS Disclosure; Fig 3; 75pp; English.
XX
The present sequence is that of amino acids 190-217 of the merozoite
surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
invention relates to the construction of recombinant human antibodies
against MSP-3, in particular MSP-3(194-257) antigen (see ABR41813). The
antibodies (see also ABR41816-21) are able to passively induce naturally
occurring non-sterile malaria immunity (termed premunition) and exert
their effect via the antibody-dependent cellular inhibition of growth
(ADCI) mechanism. They can be used in the diagnosis and treatment of
malaria. The invention also provides a method of screening a phage
library with target molecules, such as MSP-3 antigen, bound to beads,
preferably magnetic microspheres
XX
SQ Sequence 28 AA;

Query Match 92.7%; Score 115; DB 6; Length 28;
Best Local Similarity 95.8%; Pred. No. 3.4e-10;
Matches 23; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSY 24
   :|:|||||
Db 5 YERAKNAYOKANQAVLKAKEASSY 28

RESULT 4
ABR41822
ID ABR41822 standard; peptide; 31 AA.
XX
AC ABR41822;
XX
DT 11-AUG-2003 (first entry)
XX
DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
immunotherapy; vaccine; epitope.
XX
OS Plasmodium falciparum.

```

XX Key Location/Qualifiers
 FH Misc-difference 15.16
 FT /note= "linker residues"

XX GB2378949-A.
 XX 26-FEB-2003.
 XX
 XX 16-AUG-2001; 2001GB-00020057.
 XX
 XX (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.
 XX (NIEL/) NIELSEN L K.
 XX
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 XX a malarial disease.
 XX
 XX Example 3; Fig 10; 75pp; English.
 XX
 XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 XX 15 and 16 replaced by proprietary linker residues. The peptides were
 XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 XX encoded by clones obtained from a phage display system based on antibody
 XX variable region genes from the peripheral blood leucocytes of malaria
 XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 XX terminal part of the MSP-3 antigen, with highest reactivity with the
 XX middle part. RAM1 reacted only weakly with the peptides. The invention
 XX provides recombinant human antibodies against MSP-3, which are useful for
 XX the diagnosis and treatment of malaria. The recombinant human antibodies
 XX include the complementarity determining regions from the light and heavy
 XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX SQ Sequence 31 AA;
 Query Match 85.58; Score 106; DB 6; Length 31;
 Best Local Similarity 84.08; Pred. No. 8.7e-09;
 Matches 21; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 1 YERAKNAYQKANOAVLKAKKASSYD 25
 1 HERAKNAYQKANOAVLKAKKASSYD 25

RESULT 5
 ABR41823
 ID ABR41823 standard; peptide; 31 AA.
 XX
 XX ABR41823;
 XX
 XX 11-AUG-2003 (first entry)
 XX
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 XX immunotherapy; vaccine; epitope.
 XX
 XX Plasmodium falciparum.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 15.16
 FT /note= "linker residues"

XX GB2378949-A.
 XX 26-FEB-2003.
 XX
 XX 16-AUG-2001; 2001GB-00020057.
 XX
 XX (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.
 XX (NIEL/) NIELSEN L K.
 XX
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 XX a malarial disease.
 XX
 XX Example 3; Fig 10; 75pp; English.
 XX
 XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 XX 15 and 16 replaced by proprietary linker residues. The peptides were
 XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 XX encoded by clones obtained from a phage display system based on antibody
 XX variable region genes from the peripheral blood leucocytes of malaria
 XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 XX terminal part of the MSP-3 antigen, with highest reactivity with the
 XX middle part. RAM1 reacted only weakly with the peptides. The invention
 XX provides recombinant human antibodies against MSP-3, which are useful for
 XX the diagnosis and treatment of malaria. The recombinant human antibodies
 XX include the complementarity determining regions from the light and heavy
 XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX SQ Sequence 31 AA;
 Query Match 83.18; Score 103; DB 6; Length 31;
 Best Local Similarity 87.58; Pred. No. 2.5e-08;
 Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 EKAKNAYQKANOAVLKAKKASSYD 25
 1 ERAKNAYQKANOAVLKAKKASSYD 24

Db 1 ERAKNAYQKANOAVLKAKKASSYD 24

RESULT 6
 ABR41824
 ID ABR41824 standard; peptide; 31 AA.
 XX
 XX ABR41824;
 XX
 XX 11-AUG-2003 (first entry)
 XX
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 XX immunotherapy; vaccine; epitope.
 XX
 XX Plasmodium falciparum.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 15.16
 FT /note= "linker residues"

XX GB2378949-A.
 XX 26-FEB-2003.
 XX
 XX 16-AUG-2001; 2001GB-00020057.
 XX

XX Claim 4; Page 37; 52pp; French.

XX This sequence is a preferred fragment of the 48kD merozoite surface

CC antigen of P.falciparum which is useful as an immunogen to generate

CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the

CC erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN

CC field.)

XX Sequence 23 AA;

XX Query Match 77.4%; Score 96; DB 2; Length 23;

XX Best Local Similarity 84.0%; Pred. No. 2e-07; Indels 2; Gaps 1;

XX Matches 21; Conservative 2; Mismatches 0;

QY 1 YEKAKNAYQKNAQVAVLKAKKASSYD 25

DB 1 HERAKNAYQKNAQVAVLKAKKASSYD 23

RESULT 9

ABR41826

ABR41826 standard; peptide; 31 AA.

AC ABR41826;

XX 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;

XX immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15. .16

FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

PI Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3

XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of

XX a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)

XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite

XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino

XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues

XX 15 and 16 replaced by proprietary linker residues. The peptides were

XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are

XX encoded by clones obtained from a phage display system based on antibody

XX variable region genes from the peripheral blood leucocytes of malaria

XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-

XX terminal part of the MSP-3 antigen, with highest reactivity with the

XX middle part. RAM1 reacted only weakly with the peptides. The invention

XX provides recombinant human antibodies against MSP-3, which are useful for

XX the diagnosis and treatment of malaria. The recombinant human antibodies

XX include the complementarity determining regions from the light and heavy

XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

CC the diagnosis and treatment of malaria. The recombinant human antibodies

CC include the complementarity determining regions from the light and heavy

CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

QY 5 KNAYQKNAQVAVLKAKKASSYD 25

DB 1 KNAYQKNAQVAVLKAKKASSYD 21

RESULT 10

ABR41827

ID ABR41827 standard; peptide; 31 AA.

XX ABR41827;

XX 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;

XX immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15. .16

FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

PI Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3

XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of

XX a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)

XX used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite

XX surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino

XX acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues

XX 15 and 16 replaced by proprietary linker residues. The peptides were

XX examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are

XX encoded by clones obtained from a phage display system based on antibody

XX variable region genes from the peripheral blood leucocytes of malaria

XX immune human individuals. RAM2 reacted with the N-terminal, middle and C-

XX terminal part of the MSP-3 antigen, with highest reactivity with the

XX middle part. RAM1 reacted only weakly with the peptides. The invention

XX provides recombinant human antibodies against MSP-3, which are useful for

XX the diagnosis and treatment of malaria. The recombinant human antibodies

XX include the complementarity determining regions from the light and heavy

XX chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 71.0%; Score 88; DB 6; Length 31;
 Best Local Similarity 90.0%; Pred. No. 4.5e-06;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 NAYQKQAVLKAKEASSYD 25
 DB 1 NAYQKQAVLKAKEASSYD 20

RESULT 11
 ABR41828
 ID ABR41828 standard; peptide; 31 AA.

AC ABR41828;
 XX 11-AUG-2003 (first entry)

Peptide used in merozoite surface antigen MSP-3 epitope mapping.

Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 immunotherapy; vaccine; epitope.

Plasmodium falciparum.

Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"

GB2378949-A.

26-FEB-2003.

16-AUG-2001; 2001GB-00020057.

16-AUG-2001; 2001GB-00020057.

(DZIE/) DZIEGIEL M S H.

(LUND/) LUNDQUIST R.

(NIEL/) NIELSEN L K.

Dziegiel MSH, Lundquist R, Nielsen LK;

WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;

Query Match 67.7%; Score 84; DB 6; Length 31;
 Best Local Similarity 89.5%; Pred. No. 1.8e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 7 AYQKQAVLKAKEASSYD 25
 DB 1 AYQKQAVLKAKEASSYD 19

RESULT 12
 ABR41829
 ID ABR41829 standard; peptide; 31 AA.

AC ABR41829;

XX 11-AUG-2003 (first entry)

Peptide used in merozoite surface antigen MSP-3 epitope mapping.

Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 immunotherapy; vaccine; epitope.

Plasmodium falciparum.

Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"

GB2378949-A.

26-FEB-2003.

16-AUG-2001; 2001GB-00020057.

16-AUG-2001; 2001GB-00020057.

(DZIE/) DZIEGIEL M S H.

(LUND/) LUNDQUIST R.

(NIEL/) NIELSEN L K.

Dziegiel MSH, Lundquist R, Nielsen LK;

WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are encoded by clones obtained from a phage display system based on antibody variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

Sequence 31 AA;

Query Match 64.5%; Score 80; DB 6; Length 31;
 Best Local Similarity 88.9%; Pred. No. 7.2e-05;
 Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 YQKQAVLKAKEASSYD 25

DB 1 YQKQAVLKAKEASSYD 18

435

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoasicide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.
 XX

XX Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 XX

XX GB2378949-A.
 XX

XX 26-FEB-2003.
 XX

XX 16-AUG-2001; 2001GB-00020057.
 XX

XX 16-AUG-2001; 2001GB-00020057.
 XX

XX (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.
 XX (NIEL/) NIELSEN L K.
 XX

PI Dzigiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX

XX Example 3; Fig 10; 75pp; English.
 XX

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IGG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX

XX Sequence 31 AA;

Query Match 52.4%; Score 65; DB 6; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.013; Mismatches 0; Indels 0; Gaps 0;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 ANQAVLKAKEASSY 24
 |||||
 Db 1 ANQAVLKAKEASSY 14

Search completed: September 22, 2004, 18:06:19
 Job time : 41.876 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 43.0661 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-12

Perfect score: 150

Sequence: 1 AKSASYDYLGNFGGVPDHKKEN 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	150	100.0	27	2 AAR54686	Aar54686 Plasmodium
2	150	100.0	27	6 ABR41815	Abr41815 Merozoite
3	150	100.0	64	2 AAR54684	Aar54684 Plasmodium
4	150	100.0	64	6 ABR41813	Abr41813 Merozoite
5	138	92.0	31	6 ABR41835	Abr41835 Peptide u
6	137	91.3	31	6 ABR41838	Abr41838 Peptide u
7	136	90.7	31	6 ABR41839	Abr41839 Peptide u
8	134	89.3	31	6 ABR41834	Abr41834 Peptide u
9	132	88.0	31	6 ABR41840	Abr41840 Peptide u
10	131	87.3	31	6 ABR41837	Abr41837 Peptide u
11	130	86.7	31	6 ABR41836	Abr41836 Peptide u
12	127	84.7	31	6 ABR41841	Abr41841 Peptide u
13	126	84.0	31	6 ABR41833	Abr41833 Peptide u
14	124	82.7	31	6 ABR41842	Abr41842 Peptide u
15	119	79.3	31	6 ABR41832	Abr41832 Peptide u
16	118	78.7	31	6 ABR41843	Abr41843 Peptide u
17	114	76.0	31	6 ABR41831	Abr41831 Peptide u
18	113	75.3	31	6 ABR41844	Abr41844 Peptide u
19	112	74.7	31	6 ABR41830	Abr41830 Peptide u
20	109	72.7	31	6 ABR41845	Abr41845 Peptide u
21	108	72.0	31	6 ABR41829	Abr41829 Peptide u
22	103	68.7	31	6 ABR41828	Abr41828 Peptide u
23	102	68.0	31	6 ABR41846	Abr41846 Peptide u
24	99	66.0	31	6 ABR41847	Abr41847 Peptide u
25	94	62.7	31	6 ABR41827	Abr41827 Peptide u

26	92	61.3	31	6 ABR41848	Abr41848 Peptide u
27	88	58.7	31	6 ABR41849	Abr41849 Peptide u
28	88	58.7	31	6 ABR41826	Abr41826 Peptide u
29	85	56.7	31	6 ABR41850	Abr41850 Peptide u
30	85	56.7	31	6 ABR41825	Abr41825 Peptide u
31	84	56.0	31	6 ABR41851	Abr41851 Peptide u
32	84	56.0	31	6 ABR41824	Abr41824 Peptide u
33	82	54.7	31	6 ABR41823	Abr41823 Peptide u
34	79	52.7	31	6 ABR41852	Abr41852 Peptide u
35	76	50.7	31	6 ABR41822	Abr41822 Peptide u
36	74	49.3	31	6 ABR41853	Abr41853 Peptide u
37	68	45.3	31	6 ABR41854	Abr41854 Peptide u
38	65	43.3	11	6 ABR41856	Abr41856 Merozoite
39	62	41.3	31	6 ABR41855	Abr41855 Peptide u
40	58	38.7	388	6 ABM67504	Abm67504 Phototax
41	58	38.7	596	6 ABU23235	Abu23235 Protein e
42	58	38.7	599	6 ABU21874	Abu21874 Protein e
43	58	38.7	600	6 ABU20096	Abu20096 Protein e
44	58	38.7	600	6 ABU22081	Abu22081 Protein e
45	57	38.0	382	6 ABU15372	Abu15372 Protein e

ALIGNMENTS

RESULT 1

AAR54686

ID AAR54686 standard; protein; 27 AA.

XX AAR54686;

XX

DT 25-MAR-2003 (revised)

DT 15-DEC-1994 (first entry)

XX

DE Plasmodium falciparum merozoite 48kD surface protein fragment III.

XX Merozoite 48kD surface antigen; immunogen; vaccine; malaria; cytophilic antibody.

XX

OS Plasmodium falciparum.

XX

PN WO9409140-A1.

XX

PD 28-APR-1994.

XX

PF 18-OCT-1993; 93WO-FR001024.

XX

PR 19-OCT-1992; 92FR-00012488.

XX

(INSP) INST PASTEUR.

PI Druilhe P, Bouharoun-Tayoun H, Oeuvray C;

XX

DR WPI; 1994-151325/18.

XX

N-PSDB; AAQ64679.

XX

PT New Plasmodium antigen generating cytophilic - mimicking protection induced by long term exposure to the parasite, useful in vaccination against and diagnosis of malaria, and prodn of antibodies for passive immunisation.

XX

PS Claim 4; Page 37; 52pp; French.

XX

CC This sequence is a preferred fragment of the 48kD merozoite surface antigen of P.falciparum which is useful as an immunogen to generate antibodies (Abs). The Abs can induce a cytotoxic reaction against the erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 27 AA;

Query Match 100.0%; Score 150; DB 2; Length 27;

Best Local Similarity 100.0%; Pred. No. 5.7e-15;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWEGGGVPEHKKEEN 27
 DB 1 AKEASSYDYLWEGGGVPEHKKEEN 27

RESULT 2
 ABR41815

ID ABR41815 standard; protein; 27 AA.
 AC ABR41815;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Merozoite surface antigen MSP-3 peptide (aa211-237).
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 XX immunotherapy; vaccine.
 OS Plasmodium falciparum.
 XX

Key Location/Qualifiers
 FH 10..20
 FT /note= "epitope specifically referred to in Claim 3"
 XX

PN GB2378949-A.
 XX
 PD 26-FEB-2003.
 XX
 PF 16-AUG-2001; 2001GB-00020057.
 XX
 PR 16-AUG-2001; 2001GB-00020057.
 XX

(DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Dzigiel MSH, Lundquist R, Nielsen LK;
 XX
 DR WPI; 2003-303108/30.
 XX

Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX

PS Disclosure; Fig 3; 75pp; English.
 XX

The present sequence is that of amino acids 211-237 of the merozoite
 CC surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
 CC invention relates to the construction of recombinant human antibodies
 CC against MSP-3, in particular MSP-3(194-257) antigen (see ABR41813). The
 CC antibodies (see also ABR41816-21) are able to passively induce naturally
 CC occurring non-sterile malaria immunity (termed premunition) and exert
 CC their effect via the antibody-dependent cellular inhibition of growth
 CC (ADCI) mechanism. They can be used in the diagnosis and treatment of
 CC malaria. The invention also provides a method of screening a phage
 CC library with target molecules, such as MSP-3 antigen, bound to beads,
 CC preferably magnetic microspheres
 XX
 SQ Sequence 27 AA;

Query Match 100.0%; Score 150; DB 6; Length 27;
 Best Local Similarity 100.0%; Pred. No. 5.7e-15;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWEGGGVPEHKKEEN 27
 DB 1 AKEASSYDYLWEGGGVPEHKKEEN 27

RESULT 3
 AAR54684

ID AAR54684 standard; protein; 64 AA.
 AC AAR54684;
 XX
 DT 25-MAR-2003 (revised)
 DT 15-DEC-1994 (first entry)
 XX
 DE Plasmodium falciparum merozoite 48kD surface protein fragment I.
 XX
 KW Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
 KW cytophilic antibody.
 XX
 OS Plasmodium falciparum.
 XX
 PN WO9409140-A1.
 XX
 PD 28-APR-1994.
 XX
 PF 18-OCT-1993; 93WO-FR001024.
 XX
 PR 19-OCT-1992; 92FR-00012488.
 XX

(INSP) INST PASTEUR.
 PA
 PI Druilhe P, Bouharoun-Tayoun H, Oeuuvray C;
 XX
 DR WPI; 1994-151325/18.
 DR N-PSDB; AAQ64677.
 XX

New Plasmodium antigen generating cytophilic - mimicking protection
 PT induced by long term exposure to the parasite, useful in vaccination
 PT against and diagnosis of malaria, and prodn of antibodies for passive
 PT immunisation.
 XX

PS Claim 3; Page 36; 52pp; French.
 XX

This sequence is a preferred fragment of the 48kD merozoite surface
 CC antigen of P.falciparum which is useful as an immunogen to generate
 CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the
 CC erythrocyte stage of the parasite. Specific subfragments of AAR54684 are
 CC separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 64 AA;

Query Match 100.0%; Score 150; DB 2; Length 64;
 Best Local Similarity 100.0%; Pred. No. 1.4e-14;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWEGGGVPEHKKEEN 27
 DB 18 AKEASSYDYLWEGGGVPEHKKEEN 44

RESULT 4
 ABR41813

ID ABR41813 standard; protein; 64 AA.
 AC ABR41813;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Merozoite surface antigen MSP-3 peptide (aa194-257).
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine.
 XX
 OS Plasmodium falciparum.
 XX

Key Location/Qualifiers
 FH 27..37
 FT /note= "epitope specifically referred to in Claim 3"
 XX

PA	(LUND/) LUNDQUIST R.
PA	(NIEL/) NIELSEN L K.
XX	
PI	Driegiel MSH, Lundquist R, Nielsen LK;
XX	
XX	WPI; 2003-303108/30.
XX	
PT	Novel recombinant human antibody specific for merozoite surface protein-3
PT	antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT	a malarial disease.
XX	
XX	Example 3; Fig 10; 75pp; English.
XX	
CC	The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC	used in a PPSCAN epitope mapping of Plasmodium falciparum merozoite
CC	surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC	acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC	15 and 16 replaced by proprietary linker residues. The peptides were
CC	examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC	encoded by clones obtained from a phage display system based on antibody
CC	variable region genes from the peripheral blood leucocytes of malaria
CC	immune human individuals. RAM2 reacted with the N-terminal, middle and C
CC	terminal part of the MSP-3 antigen, with highest reactivity with the
CC	middle part. RAM1 reacted only weakly with the peptides. The invention
CC	provides recombinant human antibodies against MSP-3, which are useful for
CC	the diagnosis and treatment of malaria. The recombinant human antibodies
CC	include the complementarity determining regions from the light and heavy
CC	chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX	
XX	Sequence 31 AA;
XX	
Query Match	92.0%; Score 138; DB 6; Length 31;
Best Local Similarity	92.6%; Pred. No. 4e-13; 2; Indels 0; Gaps 0
Matches	25; Conservative 0; Mismatches 0
QY	1 AKEASSYDYILGWFGGVPPEHKKEEN 27
Db	5 AKEASSYDYIXXWFGGVPPEHKKEEN 31
RESULT 6	
ABR41838	
ID	ABR41838 standard; peptide; 31 AA.
XX	
AC	ABR41838;
XX	
DT	11-AUG-2003 (first entry)
XX	
DE	Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX	
KW	Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
KW	immunotherapy; vaccine; epitope.
XX	
OS	Plasmodium falciparum.
XX	
Key	Location/Qualifiers
FT	Misc-difference 15..16
XX	/note= "linker residues"
XX	
GB2378949-A.	
PN	
XX	
PD	26-FEB-2003.
XX	
XX	16-AUG-2001; 2001GB-00020057.
XX	
XX	16-AUG-2001; 2001GB-00020057.
XX	
PA	{DZIE/} DZIEGIEL M S H.
PA	(LUND/) LUNDQUIST R.
PA	(NIEL/) NIELSEN L K.
XX	
PI	Driegiel MSH, Lundquist R, Nielsen LK;
XX	

```

DR WPI; 2003-303108/30.
XX
XX Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC terminal part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
XX Sequence 31 AA;
SQ
Query Match 91.3%; Score 137; DB 6; Length 31;
Best Local Similarity 92.6%; Pred. No. 5.6e-13;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYILGWFGGVPPEHKKEEN 27
DB 2 AKEASSYDYILGWXXGGVPPEHKKEEN 28

RESULT 7
ABR41839
ID ABR41839 standard; peptide; 31 AA.
XX
XX ABR41839;
XX
XX 11-AUG-2003 (first entry)
XX
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX immunotherapy; vaccine; epitope.
XX
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH Misc-difference 15..16 /note= "linker residues"
FT
FT
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX (DZIE/) DZIEGIEL M S H.
XX (LUND/) LUNDQUIST R.
XX (NIEL/) NIELSEN L K.
XX
XX Dziegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX
XX Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC terminal part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
XX Sequence 31 AA;
SQ
Query Match 90.7%; Score 136; DB 6; Length 31;
Best Local Similarity 92.6%; Pred. No. 7.9e-13;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYILGWFGGVPPEHKKEEN 27
DB 1 AKEASSYDYILGWXXGGVPPEHKKEEN 27

RESULT 8
ABR41834
ID ABR41834 standard; peptide; 31 AA.
XX
XX ABR41834;
XX
XX 11-AUG-2003 (first entry)
XX
XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX immunotherapy; vaccine; epitope.
XX
XX Plasmodium falciparum.
XX
XX Key Location/Qualifiers
FH Misc-difference 15..16 /note= "linker residues"
FT
FT
XX
XX GB2378949-A.
XX
XX 26-FEB-2003.
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX 16-AUG-2001; 2001GB-00020057.
XX
XX (DZIE/) DZIEGIEL M S H.
XX (LUND/) LUNDQUIST R.
XX (NIEL/) NIELSEN L K.
XX
XX Dziegiel MSH, Lundquist R, Nielsen LK;
XX WPI; 2003-303108/30.
XX
XX Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.
XX
XX Example 3; Fig 10; 75pp; English.
XX
XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite

```

CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;

Query Match 89.3%; Score 134; DB 6; Length 31;
 Best Local Similarity 92.3%; Pred. No. 1.6e-12;
 Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 DB
 1 AKEASSYDYILGWFGGVPPEHKKKE 26
 6 AKEASSYDYXXGWFGGVPPEHKKKE 31

RESULT 9
 ABR41840
 ID ABR41840 standard; peptide; 31 AA.
 AC ABR41840;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX
 KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX
 OS Plasmodium falciparum.

Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"

GB2378949-A.
 26-FEB-2003.

16-AUG-2001; 2001GB-00020057.
 16-AUG-2001; 2001GB-00020057.

(DZIE/) DZIEGIEL M S H.
 (LUND/) LUNDQUIST R.
 (NIEL/) NIELSEN L K.

Dziegiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3
 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55)
 used in a pepSCAN epitope mapping of Plasmodium falciparum merozoite
 surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody

CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;

Query Match 88.0%; Score 132; DB 6; Length 31;
 Best Local Similarity 92.3%; Pred. No. 3.1e-12;
 Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEASSYDYILGWFGGVPPEHKKKEEN 27
 DB 1 KEASSYDYILGWFGGVPPEHKKKEEN 26

RESULT 10
 ABR41837
 ID ABR41837 standard; peptide; 31 AA.
 AC ABR41837;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 immunotherapy; vaccine; epitope.

Plasmodium falciparum.

Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"

GB2378949-A.
 26-FEB-2003.

16-AUG-2001; 2001GB-00020057.
 16-AUG-2001; 2001GB-00020057.

(DZIE/) DZIEGIEL M S H.
 (LUND/) LUNDQUIST R.
 (NIEL/) NIELSEN L K.

Dziegiel MSH, Lundquist R, Nielsen LK;
 WPI; 2003-303108/30.

Novel recombinant human antibody specific for merozoite surface protein-3
 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 a malarial disease.

Example 3; Fig 10; 75pp; English.

The present sequence is one of a set of 34 peptides (see ABR41822-55)
 used in a pepSCAN epitope mapping of Plasmodium falciparum merozoite
 surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for

CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX

SQ Sequence 31 AA;

Query Match 87.3%; Score 131; DB 6; Length 31;
Best Local Similarity 92.6%; Pred. No. 4.3e-12;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPPEHKKEEN 27
|||||
Db 3 AKEASSYDYLGXFXFGGVPPEHKKEEN 29
|||||

RESULT 11
ABR41836
ID ABR41836 standard; peptide; 31 AA.

XX ABR41836;

DT 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16
FT /notes= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Dzigiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC terminal part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 86.7%; Score 130; DB 6; Length 31;
Best Local Similarity 92.6%; Pred. No. 6.1e-12;
Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AKEASSYDYLGWFGGVPPEHKKEEN 27
|||||
Db 4 AKEASSYDYLXFXFGGVPPEHKKEEN 30
|||||

RESULT 12
ABR41841
ID ABR41841 standard; peptide; 31 AA.

XX ABR41841;

DT 11-AUG-2003 (first entry)

XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16
FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Dzigiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC terminal part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 84.7%; Score 127; DB 6; Length 31;
Best Local Similarity 92.0%; Pred. No. 1.7e-11;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EASSYDYLWFGGVPPEHKKEN 27
 Db 1 EASSYDYLWFGGVPPEHKKEN 25

RESULT 13

ABR41833
 ID ABR41833 standard; peptide; 31 AA.

XX AC ABR41833;

XX 11-AUG-2003 (first entry)

Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers
 FT Misc-difference 15..16
 FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

PA (LUND/) LUNDQUIST R.

PA (NIEL/) NIELSEN L K.

XX Driegiel MSH; Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 84.0%; Score 126; DB 6; Length 31;
 Best Local Similarity 92.0%; Pred. No. 2.4e-11;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AKASSYDYLWFGGVPPEHKKKE 25

Db 7 AKASSYDYLWFGGVPPEHKKKE 31

RESULT 14

ABR41842
 ID ABR41842 standard; peptide; 31 AA.

XX ABR41842;

XX 11-AUG-2003 (first entry)

Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

XX Plasmodium falciparum.

XX Key Location/Qualifiers

FT Misc-difference 15..16
 FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

PA (LUND/) LUNDQUIST R.

PA (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 82.7%; Score 124; DB 6; Length 31;
 Best Local Similarity 91.7%; Pred. No. 4.7e-11;
 Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 ASSYDYLWFGGVPPEHKKEN 27

Db 1 ASSYDYLWFGGVPPEHKKEN 24

RESULT 15

ABR41832

ID ABR41832 standard; peptide; 31 AA.

XX ABR41832;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:00:42 ; Search time 44.6612 Seconds
(without alignments)
177.141 Million cell updates/sec

Title: US-10-774-602-13

Perfect score: 147

Sequence: 1 PEHKKENMLSHLYVSSKDKENISKEND 28

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	97.3	28	2 AAR54687	Aar54687 Plasmodiu
2	143	97.3	64	2 AAR54684	Aar54684 Plasmodiu
3	143	97.3	64	6 ABR41813	Abr41813 Merozoite
4	129	87.8	31	6 ABR41855	Abr41855 Peptide u
5	128	87.1	31	6 ABR41854	Abr41854 Peptide u
6	126	85.7	31	6 ABR41853	Abr41853 Peptide u
7	119	81.0	31	6 ABR41852	Abr41852 Peptide u
8	112	76.2	31	6 ABR41851	Abr41851 Peptide u
9	108	73.5	31	6 ABR41850	Abr41850 Peptide u
10	105	71.4	31	6 ABR41849	Abr41849 Peptide u
11	99	67.3	31	6 ABR41848	Abr41848 Peptide u
12	94	63.9	31	6 ABR41847	Abr41847 Peptide u
13	86	58.5	31	6 ABR41846	Abr41846 Peptide u
14	82	55.8	31	6 ABR41842	Abr41842 Peptide u
15	80	54.4	31	6 ABR41845	Abr41845 Peptide u
16	79	53.7	31	6 ABR41843	Abr41843 Peptide u
17	78	53.1	31	6 ABR41844	Abr41844 Peptide u
18	78	53.1	31	6 ABR41841	Abr41841 Peptide u
19	71	48.3	31	6 ABR41840	Abr41840 Peptide u
20	67	45.6	31	6 ABR41839	Abr41839 Peptide u
21	59	40.1	31	6 ABR41838	Abr41838 Peptide u
22	55	37.4	31	6 ABR41837	Abr41837 Peptide u
23	53	36.1	353	4 AAM41451	Aam41451 Human pol
24	53	36.1	426	3 AAB42986	Aab42986 Human ORF
25	53	36.1	426	4 AAM39665	Aam39665 Human pol

26	53	36.1	426	4 AAB94867	Aab94867 Human pro
27	53	36.1	426	4 AAG89197	Aag89197 Human sec
28	53	36.1	426	4 AAG89190	Aag89190 Human sec
29	53	36.1	426	7 ADC06747	Adc06747 Human ino
30	53	36.1	481	4 AAG89295	Aag89295 Human sec
31	53	36.1	3242	4 ABB58857	Abb58857 Drosophil
32	52	35.4	443	6 ABU20833	Abu20833 Protein e
33	52	35.4	693	6 ADA14410	Ada14410 Mouse spe
34	52	35.4	2206	3 AAB18254	Aab18254 Plasmodiu
35	51.5	35.0	407	3 AAB18248	Aab18248 Plasmodiu
36	51.5	35.0	1182	6 ABU25521	Abu25521 Protein e
37	51	34.7	31	6 ABR41836	Abr41836 Peptide u
38	51	34.7	174	4 ABB52591	Abb52591 Escherich
39	51	34.7	415	6 ABU42518	Abu42518 Protein e
40	50.5	34.4	296	4 ABB69513	Abb69513 Drosophil
41	50	34.0	218	5 ABP08273	Abp08273 Human ORF
42	50	34.0	518	6 ABU25405	Abu25405 Protein e
43	50	34.0	1997	3 AAB18287	Aab18287 Plasmodiu
44	49.5	33.7	774	6 ABM73172	Abm73172 Staphyloc
45	49.5	33.7	1714	3 AAB18275	Aab18275 Plasmodiu

ALIGNMENTS

RESULT 1

AAR54687

ID AAR54687 standard; protein; 28 AA.

XX AAR54687;

XX 25-MAR-2003 (revised)

DT 15-DEC-1994 (first entry)

XX Plasmodium falciparum merozoite 48kD surface protein fragment IV.

DE Plasmodium falciparum merozoite 48kD surface protein fragment IV.

XX Merozoite 48kD surface antigen; immunogen; vaccine; malaria;

KW cytophilic antibody.

XX Plasmodium falciparum.

OS Plasmodium falciparum.

XX WO9409140-A1.

PN 28-APR-1994.

XX 18-OCT-1993; 93WO-FR001024.

PF 19-OCT-1992; 92FR-00012488.

XX (INSP) INST PASTEUR.

PI Druilhe P, Bouharoun-Tayoun H, Oeuvery C;

XX WPI; 1994-151325/18.

DR N-PSDB; AAQ64680.

PT New Plasmodium antigen generating cytophilic - mimicking protection induced by long term exposure to the parasite, useful in vaccination against and diagnosis of malaria, and prodn of antibodies for passive immunisation.

XX Claim 4; Page 37; 52pp; French.

CC This sequence is a preferred fragment of the 48kD merozoite surface antigen of P.falciparum which is useful as an immunogen to generate antibodies (Abs). The Abs can induce a cytotoxic reaction against the erythrocyte stage of the parasite. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 28 AA;

SQ

Query Match 97.3%; Score 143; DB 2; Length 28;
Best Local Similarity 96.4%; Pred. No. 4.5e-13;

Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 1 PEHKKEENMLSHLYVSSKDKENISKENE 28

RESULT 2
 AAR54684
 ID AAR54684 standard; protein; 64 AA.
 XX AAR54684;
 XX
 DT 25-MAR-2003 (revised)
 DT 15-DEC-1994 (first entry)
 XX
 XX Plasmodium falciparum merozoite 48kD surface protein fragment I.
 DE Merozoite 48kD surface antigen; immunogen; vaccine; malaria;
 DE cytophilic antibody.
 OS Plasmodium falciparum.
 XX WO9409140-A1.
 XX
 PN 28-APR-1994.
 PD
 XX 18-OCT-1993; 93WO-FR001024.
 PF
 XX 19-OCT-1992; 92FR-00012488.
 PR
 XX (INSP) INST PASTEUR.
 PA
 XX Druiilhe P, Bouharoun-Tayoun H, Oeuvray C;
 PI WPI; 1994-151325/18.
 XX N-PSDB; AAQ64677.
 DR
 XX New Plasmodium antigen generating cytophilic - mimicking protection
 PT induced by long term exposure to the parasite, useful in vaccination
 PT against and diagnosis of malaria, and prodn of antibodies for passive
 PT immunisation.
 XX
 PS Claim 3; Page 36; 52pp; French.
 XX
 XX This sequence is a preferred fragment of the 48kD merozoite surface
 CC antigen of P.falciparum which is useful as an immunogen to generate
 CC antibodies (Abs). The Abs can induce a cytotoxic reaction against the
 CC erythrocyte stage of the parasite. Specific subfragments of AAR54684 are
 CC separately claimed (see AAR54685-R54687). (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 64 AA;

Query Match 97.3%; Score 143; DB 2; Length 64;
 Best Local Similarity 96.4%; Pred. No. 1.2e-12;
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 37 PEHKKEENMLSHLYVSSKDKENISKENE 64

RESULT 3
 ABR41813
 ID ABR41813 standard; protein; 64 AA.
 XX ABR41813;
 AC
 XX 11-AUG-2003 (first entry)
 DT
 XX Merozoite surface antigen MSP-3 peptide (aa194-257).
 DE
 XX

Query Match 97.3%; Score 143; DB 2; Length 64;
 Best Local Similarity 96.4%; Pred. No. 1.2e-12;
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 37 PEHKKEENMLSHLYVSSKDKENISKENE 64

RESULT 4
 ABR41855
 ID ABR41855 standard; peptide; 31 AA.
 XX
 AC ABR41855;
 DT
 XX 11-AUG-2003 (first entry)
 DT
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 KW Plasmodium falciparum.
 XX
 OS Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine.
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 FH 27..37
 FT /note= "epitope specifically referred to in Claim 3"
 FT
 XX GB2378949-A.
 PN
 XX 26-FEB-2003.
 PD
 XX 16-AUG-2001; 2001GB-00020057.
 PF
 XX 16-AUG-2001; 2001GB-00020057.
 PR
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX
 PI Dziegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 DR
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 PT
 XX Disclosure; Fig 3; 75pp; English.
 PS
 XX The present sequence is that of amino acids 194-257 of the merozoite
 CC surface protein 3 (MSP-3) of Plasmodium falciparum clone D10. The
 CC invention relates to the construction of recombinant human antibodies
 CC against MSP-3, in particular MSP-3(194-257) antigen. The antibodies (see
 CC also ABR41816-21) are able to passively induce naturally occurring non-
 CC sterile malaria immunity (termed premunition) and exert their effect via
 CC the antibody-dependent cellular inhibition of growth (ADCI) mechanism.
 CC They can be used in the diagnosis and treatment of malaria. The invention
 CC also provides a method of screening a phage library with target
 CC molecules, such as MSP-3 antigen, bound to beads, preferably magnetic
 CC microspheres
 XX
 SQ Sequence 64 AA;

Query Match 97.3%; Score 143; DB 6; Length 64;
 Best Local Similarity 96.4%; Pred. No. 1.2e-12;
 Matches 27; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
 DB 37 PEHKKEENMLSHLYVSSKDKENISKENE 64

RESULT 4
 ABR41855
 ID ABR41855 standard; peptide; 31 AA.
 XX
 AC ABR41855;
 DT
 XX 11-AUG-2003 (first entry)
 DT
 XX Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 KW Plasmodium falciparum.
 XX
 OS Key Location/Qualifiers
 FH Misc-difference 15..16
 FT /note= "linker residues"
 FT

PN GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 10; 75pp; English.
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IGG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX Sequence 31 AA;
 SQ
 Query Match 87.8%; Score 129; DB 6; Length 31;
 Best Local Similarity 89.3%; Pred. No. 4.8e-11;
 Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PEHKKENMLSHLYVSSKDKENISKEND 28
 DB 4 PEHKKENMLSHLYVSSKDKENISKENE 31
 ID 1854
 AC ABR41854 standard; peptide; 31 AA.
 XX ABR41854;
 XX 11-AUG-2003 (first entry)
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 PH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 XX WPI; 2003-303108/30.
 XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 10; 75pp; English.
 CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IGG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX Sequence 31 AA;
 SQ
 Query Match 87.8%; Score 129; DB 6; Length 31;
 Best Local Similarity 89.3%; Pred. No. 4.8e-11;
 Matches 25; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PEHKKENMLSHLYVSSKDKENISKEND 28
 DB 4 PEHKKENMLSHLYVSSKDKENISKENE 31
 ID 1854
 AC ABR41854 standard; peptide; 31 AA.
 XX ABR41854;
 XX 11-AUG-2003 (first entry)
 DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 XX Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.
 XX Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 PH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX (DZIE/) DZIEGIEL M S H.
 PA (LUND/) LUNDQUIST R.

XX 16-AUG-2001; 2001GB-00020057.
 PR (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.
 PA (NIEL/) NIELSEN L K.
 XX Driegiel MSH, Lundquist R, Nielsen LK;
 PI WPI; 2003-303108/30.
 DR Novel recombinant human antibody specific for merozoite surface protein-3
 XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.
 XX Example 3; Fig 10; 75pp; English.
 PS The present sequence is one of a set of 34 peptides (see ABR41822-55)
 XX used in a PEPCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IGG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX Sequence 31 AA;
 SQ
 Query Match 87.1%; Score 128; DB 6; Length 31;
 Best Local Similarity 92.6%; Pred. No. 6.6e-11;
 Matches 25; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PEHKKENMLSHLYVSSKDKENISKEN 27
 DB 5 PEHKKENMLSHLYVSSKDKENISKEN 31
 RESULT 6
 ABR41853
 ID ABR41853 standard; peptide; 31 AA.
 XX ABR41853;
 AC ABR41853;
 XX 11-AUG-2003 (first entry)
 DT Peptide used in merozoite surface antigen MSP-3 epitope mapping.
 DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 XX immunotherapy; vaccine; epitope.
 KW Plasmodium falciparum.
 OS
 XX Key Location/Qualifiers
 PH Misc-difference 15..16
 FT /note= "linker residues"
 FT
 XX GB2378949-A.
 XX 26-FEB-2003.
 XX 16-AUG-2001; 2001GB-00020057.
 XX 16-AUG-2001; 2001GB-00020057.
 PR (DZIE/) DZIEGIEL M S H.
 XX (LUND/) LUNDQUIST R.

PS Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

SQ Query Match 76.2%; Score 112; DB 6; Length 31; Best Local Similarity 91.7%; Pred. No. 1.2e-08; Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PEHKEENMLSHLYVSSKDKENIS 24
Db 8 PEHKEEXXLHLYVSSKDKENIS 31

RESULT 9

ABR41850

ID ABR41850 standard; peptide; 31 AA.

XX AC ABR41850;

XX DT 11-AUG-2003 (first entry)

XX DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide; immunotherapy; vaccine; epitope.

XX OS Plasmodium falciparum.

XX FH Key Location/Qualifiers

FT Misc-difference 15..16 /note= "linker residues"

FT GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

PA (LUND/) LUNDQUIST R.

PA (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

XX Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

SQ Query Match 73.5%; Score 108; DB 6; Length 31; Best Local Similarity 91.3%; Pred. No. 4.3e-08; Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PEHKEENMLSHLYVSSKDKENI 23
Db 9 PEHKEEXXLHLYVSSKDKENI 31

RESULT 10

ABR41849

ID ABR41849 standard; peptide; 31 AA.

XX AC ABR41849;

XX DT 11-AUG-2003 (first entry)

XX DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

XX DE Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide; immunotherapy; vaccine; epitope.

XX OS Plasmodium falciparum.

XX FH Key Location/Qualifiers

FT Misc-difference 15..16 /note= "linker residues"

FT GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

PA (LUND/) LUNDQUIST R.

PA (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3 antigen of Plasmodium falciparum, useful for treatment and prophylaxis of a malarial disease.

XX Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55) used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues 15 and 16 replaced by proprietary linker residues. The peptides were examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are variable region genes from the peripheral blood leucocytes of malaria immune human individuals. RAM2 reacted with the N-terminal, middle and C-terminal part of the MSP-3 antigen, with highest reactivity with the middle part. RAM1 reacted only weakly with the peptides. The invention provides recombinant human antibodies against MSP-3, which are useful for the diagnosis and treatment of malaria. The recombinant human antibodies include the complementarity determining regions from the light and heavy chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;

Query Match 71.4%; Score 105; DB 6; Length 31;
 Best Local Similarity 90.9%; Pred. NO. 1.1e-07;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKOKEN 22
 DB 10 PEHKXXENMLSHLYVSSKOKEN 31

DT 11

ABR41848
 ID ABR41848 standard; peptide; 31 AA.

AC ABR41848;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

OS Plasmodium falciparum.

FH Key Location/Qualifiers
 FT Misc-difference 15..16 /note= "linker residues"

XX GB2378949-A.

PD 26-FEB-2003.

PF 16-AUG-2001; 2001GB-00020057.

PR 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

PI Dzigiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies

CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
 XX
 SQ Sequence 31 AA;

Query Match 67.3%; Score 99; DB 6; Length 31;
 Best Local Similarity 90.5%; Pred. NO. 8e-07;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDXE 21
 DB 11 PEHKXXENMLSHLYVSSKDXE 31

RESULT 12

ABR41847

ID ABR41847 standard; peptide; 31 AA.

XX ABR41847;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
 KW immunotherapy; vaccine; epitope.

OS Plasmodium falciparum.

FH Key Location/Qualifiers

FT Misc-difference 15..16 /note= "linker residues"

XX GB2378949-A.

PD 26-FEB-2003.

PF 16-AUG-2001; 2001GB-00020057.

PR 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

PI Dzigiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3
 PT antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
 PT a malarial disease.

XX Example 3; Fig 10; 75pp; English.

XX The present sequence is one of a set of 34 peptides (see ABR41822-55)
 CC used in a PEPSCAN epitope mapping of Plasmodium falciparum merozoite
 CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
 CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
 CC 15 and 16 replaced by proprietary linker residues. The peptides were
 CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
 CC encoded by clones obtained from a phage display system based on antibody
 CC variable region genes from the peripheral blood leucocytes of malaria
 CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
 CC terminal part of the MSP-3 antigen, with highest reactivity with the
 CC middle part. RAM1 reacted only weakly with the peptides. The invention
 CC provides recombinant human antibodies against MSP-3, which are useful for
 CC the diagnosis and treatment of malaria. The recombinant human antibodies
 CC include the complementarity determining regions from the light and heavy
 CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 63.9%; Score 94; DB 6; Length 31;
Best Local Similarity 90.0%; Pred. No. 4.1e-06;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PEKKKEENMLSHLYVSSKDK 20
DB 12 PEKKKEENMLSHLYVSSKDK 31

RESULT 13
ABR41846
ID ABR41846 standard; peptide; 31 AA.

XX ABR41846;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
immunotherapy; vaccine; epitope.

OS Plasmodium falciparum.

XX Key Location/Qualifiers
FH Misc-difference 15. .16
FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3

XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of

XX a malarial disease.

XX Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC middle part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 58.5%; Score 86; DB 6; Length 31;
Best Local Similarity 89.5%; Pred. No. 5.4e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PEKKKEENMLSHLYVSSKDK 19

DB 13 PEKKKEENMLSHLYVSSKDK 31

RESULT 14
ABR41842

ID ABR41842 standard; peptide; 31 AA.

XX ABR41842;

DT 11-AUG-2003 (first entry)

DE Peptide used in merozoite surface antigen MSP-3 epitope mapping.

KW Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
immunotherapy; vaccine; epitope.

OS Plasmodium falciparum.

XX Key Location/Qualifiers
FH Misc-difference 15. .16
FT /note= "linker residues"

XX GB2378949-A.

XX 26-FEB-2003.

XX 16-AUG-2001; 2001GB-00020057.

XX 16-AUG-2001; 2001GB-00020057.

XX (DZIE/) DZIEGIEL M S H.

XX (LUND/) LUNDQUIST R.

XX (NIEL/) NIELSEN L K.

XX Driegiel MSH, Lundquist R, Nielsen LK;

XX WPI; 2003-303108/30.

XX Novel recombinant human antibody specific for merozoite surface protein-3

XX antigen of Plasmodium falciparum, useful for treatment and prophylaxis of

XX a malarial disease.

XX Example 3; Fig 10; 75pp; English.

CC The present sequence is one of a set of 34 peptides (see ABR41822-55)
CC used in a PEPCAN epitope mapping of Plasmodium falciparum merozoite
CC surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
CC acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
CC 15 and 16 replaced by proprietary linker residues. The peptides were
CC examined for reactivity with RAM1 and RAM2 IgG1. RAM1 and RAM2 are
CC encoded by clones obtained from a phage display system based on antibody
CC variable region genes from the peripheral blood leucocytes of malaria
CC immune human individuals. RAM2 reacted with the N-terminal, middle and C-
CC middle part of the MSP-3 antigen, with highest reactivity with the
CC middle part. RAM1 reacted only weakly with the peptides. The invention
CC provides recombinant human antibodies against MSP-3, which are useful for
CC the diagnosis and treatment of malaria. The recombinant human antibodies
CC include the complementarity determining regions from the light and heavy
CC chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)

XX Sequence 31 AA;

Query Match 55.8%; Score 82; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PEKKKEENMLSHLYV 15

DB 17 PEKKKEENMLSHLYV 31

RESULT 15

```

ABR41845
ID  ABR41845 standard; peptide; 31 AA.
XX
AC  ABR41845;
XX
DT  11-AUG-2003 (first entry)
XX
DE  Peptide used in merozoite surface antigen MSP-3 epitope mapping.
XX
KW  Merozoite surface protein; MSP-3; malaria; antimalarial; protozoacide;
XX  immunotherapy; vaccine; epitope.
XX
OS  Plasmodium falciparum.
XX
XX  Key      Location/Qualifiers
FH  Misc-difference 15..16
FT  /note= "linker residues"
XX
XX  GB2378949-A.
XX
XX  26-FEB-2003.
XX
XX  16-AUG-2001; 2001GB-00020057.
XX
XX  16-AUG-2001; 2001GB-00020057.
XX
XX  (DZIE/) DZIEGIEL M S H.
XX  (LUND/) LUNDQUIST R.
XX  (NIEL/) NIELSEN L K.
XX
XX  Dziegiel MSH, Lundquist R, Nielsen LK;
XX  WPI; 2003-303108/30.
XX
XX  Novel recombinant human antibody specific for merozoite surface protein-3
XX  antigen of Plasmodium falciparum, useful for treatment and prophylaxis of
XX  a malarial disease.
XX
XX  Example 3; Fig 10; 75pp; English.
XX
XX  The present sequence is one of a set of 34 peptides (see ABR41822-55)
XX  used in a pepSCAN epitope mapping of Plasmodium falciparum merozoite
XX  surface antigen 3 (MSP-3). Each peptide covers a stretch of 31 amino
XX  acids within amino acids 194-257 (see ABR41813) of MSP-3, with residues
XX  15 and 16 replaced by proprietary linker residues. The peptides were
XX  examined for reactivity with RAM1 and RAM2 IGI1. RAM1 and RAM2 are
XX  encoded by clones obtained from a phage display system based on antibody
XX  variable region genes from the peripheral blood leucocytes of malaria
XX  immune human individuals. RAM2 reacted with the N-terminal, middle and C-
XX  terminal part of the MSP-3 antigen, with highest reactivity with the
XX  middle part. RAM1 reacted only weakly with the peptides. The invention
XX  provides recombinant human antibodies against MSP-3, which are useful for
XX  the diagnosis and treatment of malaria. The recombinant human antibodies
XX  include the complementarity determining regions from the light and heavy
XX  chain variable regions of RAM1, RAM2 and RAM3 (see ABR41816-21)
XX
XX  Query Match      54.4%; Score 80; DB 6; Length 31;
XX  Best Local Similarity 88.9%; Pred. No. 0.00038;
XX  Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX  QY  1 PEHKKENMLSHLYVSSK 18
XX      | | | | | | | | | |
XX  Db  14 PXXKKENMLSHLYVSSK 31

```

Search completed: September 22, 2004, 18:06:21
Job time : 45.6612 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 26.4463 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-11

Perfect score: 124

Sequence: 1 YEKAKYQKANOAVLKAEASSYD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	124	100.0	354	5 Q25995	Q25995 plasmodium
2	124	100.0	354	5 Q81J55	Q81J55 plasmodium
3	124	100.0	380	5 Q26019	Q26019 plasmodium
4	121	97.6	361	5 Q95P15	Q95P15 plasmodium
5	121	97.6	379	5 Q25705	Q25705 plasmodium
6	121	97.6	379	5 Q25706	Q25706 plasmodium
7	121	97.6	379	5 Q9UC64	Q9UC64 plasmodium
8	120	96.8	329	5 Q9NFV9	Q9NFV9 plasmodium
9	119	96.0	346	5 Q9U0G0	Q9U0G0 plasmodium
10	118	95.2	157	5 Q15854	Q15854 plasmodium
11	118	95.2	157	5 Q15855	Q15855 plasmodium
12	118	95.2	157	5 Q9TV90	Q9TV90 plasmodium
13	118	95.2	157	5 Q15856	Q15856 plasmodium
14	118	95.2	157	5 Q15853	Q15853 plasmodium
15	118	95.2	157	5 Q9TV91	Q9TV91 plasmodium
16	116	93.5	64	5 Q15762	Q15762 plasmodium

17	115	92.7	183	5 Q9TY89	Q9TY89 plasmodium
18	115	92.7	183	5 Q15857	Q15857 plasmodium
19	115	92.7	183	5 Q9TY87	Q9TY87 plasmodium
20	115	92.7	183	5 Q15858	Q15858 plasmodium
21	115	92.7	183	5 Q9TY88	Q9TY88 plasmodium
22	115	92.7	183	5 Q9TY86	Q9TY86 plasmodium
23	115	92.7	183	5 Q15859	Q15859 plasmodium
24	115	92.7	183	5 Q15862	Q15862 plasmodium
25	111	89.5	190	5 Q15860	Q15860 plasmodium
26	110	88.7	183	5 Q15863	Q15863 plasmodium
27	108	87.1	180	5 Q9NIG7	Q9NIG7 plasmodium
28	102	82.3	183	5 Q15861	Q15861 plasmodium
29	59	47.6	890	2 Q01891	Q01891 enterococcu
30	59	47.6	890	16 Q82YS1	Q82YS1 enterococcu
31	57	46.0	639	5 Q86RM4	Q86RM4 plasmodium
32	55.5	44.8	652	10 Q9ASS0	Q9ASS0 arabidopsis
33	55.5	44.8	820	10 Q9LFE0	Q9LFE0 arabidopsis
34	55	44.4	752	16 Q9CGC9	Q9CGC9 lactococcus
35	55	44.4	891	16 Q04111	Q04111 enterococcu
36	54.5	44.0	1037	3 Q9P3K1	Q9P3K1 neutrospora
37	52	41.9	213	5 P90899	P90899 caenorhabdi
38	52	41.9	231	5 Q86D10	Q86D10 caenorhabdi
39	52	41.9	420	10 Q9LZL9	Q9LZL9 arabidopsis
40	51.5	41.5	647	10 Q7X7G6	Q7X7G6 oryza sativ
41	51	41.1	169	5 Q9NIG6	Q9NIG6 plasmodium
42	51	41.1	420	10 Q8LFK4	Q8LFK4 arabidopsis
43	51	41.1	742	2 Q05082	Q05082 streptococc
44	51	41.1	901	16 Q839L8	Q839L8 enterococcu
45	51	41.1	972	2 Q8KYV2	Q8KYV2 uncultured

ALIGNMENTS

RESULT 1

Q25995	ID	Q25995	PRELIMINARY;	PRT;	354 AA.
AC	Q25995;				
DT	01-NOV-1996	(TrEMBLrel. 01, Created)			
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)			
DT	01-OCT-2002	(TrEMBLrel. 22, Last annotation update)			
DE	Antigen.				
OS	Plasmodium falciparum (isolate NF54).				
OC	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.				
OX	NCBI_TaxID=5843;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NF54;				
RX	MEDLINE=95198774; PubMed=7891748;				
RA	McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M., Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;				
RA	"Molecular variation in a novel polymorphic antigen associated with Plasmodium falciparum merozoites.";				
RT	Mol. Biochem. Parasitol. 90:21-31(1997).				
RL	EMBL; L28825; AAC09377.1; .				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=NF54;				
RX	MEDLINE=98156743; PubMed=9497029;				
RA	McColl D.J., Anders R.F.;				
RT	"Conservation of structural motifs and antigenic diversity in the Plasmodium falciparum merozoite surface protein-3 (MSP-3).";				
RL	Mol. Biochem. Parasitol. 90:21-31(1997).				
DR	EMBL; L28825; AAC09377.1; .				
SQ	SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;				

Query Match 100.0%; Score 124; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKYQKANOAVLKAEASSYD 25

Db 167 YEKAKYQKANOAVLKAEASSYD 191

```
RESULT 2
Q81J55
ID Q81J55 PRELIMINARY; PRT; 354 AA.
AC Q81J55;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Merozoite surface protein 3.
GN PF10_0345.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
  Carlson J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
  Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
  Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
  Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
  Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
  McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
  Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
  Fraser C.M., Barrell B.,
  "Genome sequence of the human malaria parasite Plasmodium
  falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014834; AAN35542.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 124; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.1e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 167 YEKAKNAYOKANQAVLKAKEASSYD 191

RESULT 3
Q26019
ID Q26019 PRELIMINARY; PRT; 380 AA.
AC Q26019;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=95198774; PubMed=7891748;
RA McCall D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
  Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
  "Molecular variation in a novel polymorphic antigen associated with
  Plasmodium falciparum merozoites."
RL Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=98156743; PubMed=9497029;
RA McCall D.J., Anders R.F.;
  "Conservation of structural motifs and antigenic diversity in the
  Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L07944; AAC09378.1; -.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 380 POLYMORPHIC ANTIGEN.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 124; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 194 YEKAKNAYOKANQAVLKAKEASSYD 218

RESULT 4
Q95P15
ID Q95P15 PRELIMINARY; PRT; 361 AA.
AC Q95P15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Hisaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
  "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus
  RT Monkeys."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY044180; AAK94780.1; -.
FT NON_TER 1 361
FT NON_TER 1 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 361;
Best Local Similarity 96.0%; Pred. No. 1.3e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 173 YEKAKNAYOKANQAVLKAKEASSYD 197

RESULT 5
Q25705
ID Q25705 PRELIMINARY; PRT; 379 AA.
AC Q25705;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X1;
RX MEDLINE=98156743; PubMed=9497029;
RA McCall D.J., Anders R.F.;
  "Conservation of structural motifs and antigenic diversity in the
  Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08851; AAC47831.1; -.
SQ SEQUENCE 379 AA; 43344 MW; DC7AF106887C9AA0 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 379;
Best Local Similarity 96.0%; Pred. No. 1.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYOKANQAVLKAKEASSYD 25
Db 193 YEKAKNAYOKANQAVLKAKEASSYD 217
```

```
RESULT 6
Q25706 PRELIMINARY; PRT; 379 AA.
AC Q25706;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2002 (TrEMBLrel. 01, Last sequence update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=csl2;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; U08852; AAC47832.1; -.
SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 379;
Best Local Similarity 96.0%; Pred. No. 1.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 193 YERAKNAYQKANOAVLKAKEASSYD 217

RESULT 7
Q9U6C4 PRELIMINARY; PRT; 379 AA.
AC Q9U6C4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
GN MSP-3.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FC1/HN;
RX Li X.R., Yu X.B., Luo S.H., Shan Z.X., Fang J.M., Xu J.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188190; AAF04099.1; -.
SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;

Query Match 97.6%; Score 121; DB 5; Length 379;
Best Local Similarity 96.0%; Pred. No. 1.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 193 YERAKNAYQKANOAVLKAKEASSYD 217

RESULT 8
Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.

OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252287; CAB85901.1; -.
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; 36916 MW; C5B045DB5E21A159 CRC64;

Query Match 96.8%; Score 120; DB 5; Length 329;
Best Local Similarity 96.0%; Pred. No. 1.7e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 157 YERAKNAYQKANOAVLKAKEASSYD 181

RESULT 9
Q9U0G0 PRELIMINARY; PRT; 346 AA.
AC Q9U0G0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252286; CAB65754.1; -.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;

Query Match 96.0%; Score 119; DB 5; Length 346;
Best Local Similarity 96.0%; Pred. No. 2.4e-08;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANOAVLKAKEASSYD 25
Db 170 YERAKNAYQKANOAVLKAKEASSYD 194

RESULT 10
O15854 PRELIMINARY; PRT; 157 AA.
AC O15854;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
```

RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RT surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234(1997).
DR EMBL; AF001138; AAC47663.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17009 MW; PDF17D10C45E74DB CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANKQAVLKAKEASSY 24
|||||
Db 134 YEKAKNAYQKANKQAVLKAKEASSY 157

T 11
C 5
ID O15855 PRELIMINARY; PRT; 157 AA.
AC O15855;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RT surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234(1997).
DR EMBL; AF001140; AAC47665.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 16981 MW; E2FE6D10C44963DB CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANKQAVLKAKEASSY 24
|||||
Db 134 YEKAKNAYQKANKQAVLKAKEASSY 157

RESULT 12
Q9TY90
ID Q9TY90 PRELIMINARY; PRT; 157 AA.
AC Q9TY90;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RT surface antigen 3.";
RL Mol. Biochem. Parasitol. 87:231-234(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=01/3;

RA Eisen D.P., Saul A., Fryauff D.J., Reeder J.C., Coppel R.L.;
RT "Extended antigenic haplotyping of Plasmodium falciparum field
RT isolates shows marked complexity.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001142; AAC47667.1; -.
DR EMBL; AF213689; AAC59916.1; -.
KW Merozoite.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17007 MW; F06E6080DB4963DB CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANKQAVLKAKEASSY 24
|||||
Db 134 YEKAKNAYQKANKQAVLKAKEASSY 157

RESULT 13
O15856
ID O15856 PRELIMINARY; PRT; 157 AA.
AC O15856;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RT surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234(1997).
DR EMBL; AF001141; AAC47666.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17034 MW; F12B7B817053D270 CRC64;

Query Match 95.2%; Score 118; DB 5; Length 157;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YEKAKNAYQKANKQAVLKAKEASSY 24
|||||
Db 134 YEKAKNAYQKANKQAVLKAKEASSY 157

RESULT 14
O15853
ID O15853 PRELIMINARY; PRT; 157 AA.
AC O15853;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN SPAM.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RT surface protein 3.";
RL Mol. Biochem. Parasitol. 87:231-234(1997).
DR EMBL; AF001137; AAC47662.1; -.
RN [2]

```

FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 16981 MW; E2FE6D10C4963DB CRC64;

Query Match
Best Local Similarity 95.2%; Score 118; DB 5; Length 157;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKKEASSY 24
   |||||
Db 134 YEKAKNAYQKANOAVLKAKKEASSY 157

RESULT 15
Q9TY91
ID Q9TY91 PRELIMINARY; PRT; 157 AA.
AC Q9TY91;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
   Merozoite surface protein 3 (Fragment).
   SPAM.
OC Plasmodium falciparum.
OX Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5633;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97391128; PubMed=9247935;
RA Huber W., Felger I., Matile H., Lipps H.J., Steiger S., Beck H.P.;
RT "Limited sequence polymorphism in the Plasmodium falciparum merozoite
RL surface protein 3."
RL Mol. Biochem. Parasitol. 87:231-234 (1997).
DR EMBL; AF001139; AAC47664.1; -.
FT NON_TER 1 1
FT NON_TER 157 157
SQ SEQUENCE 157 AA; 17007 MW; F06E6080DB4963DB CRC64;

Query Match
Best Local Similarity 95.2%; Score 118; DB 5; Length 157;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 YEKAKNAYQKANOAVLKAKKEASSY 24
   |||||
Db 134 YEKAKNAYQKANOAVLKAKKEASSY 157

```

Search completed: September 22, 2004, 18:09:05
 Job time : 29.4463 secs

100-100000-100000

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 28.562 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-12

Perfect score: 150
Sequence: 1 AKASDYVILGWFGGVPEHKKEEN 27

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

1 number of hits satisfying chosen parameters: 1017041

Maximum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	100.0	64	5	015762 plasmodium
2	150	100.0	329	5	Q9NFV9
3	150	100.0	354	5	Q25995
4	150	100.0	354	5	Q8IJ55
5	150	100.0	361	5	Q95P15
6	150	100.0	379	5	Q25705
7	150	100.0	379	5	Q25706
8	150	100.0	379	5	Q9U6C4
9	150	100.0	380	5	Q26019
10	145	96.7	346	5	Q9U0G0
11	66.5	44.3	371	5	Q9GS22
12	66.5	44.3	371	5	Q8IJ54
13	66	44.0	405	5	Q8IJ48
14	61	40.7	1238	16	O69947
15	58	38.7	596	16	Q7WR42
16	58	38.7	596	16	Q7W275

```

17 58 38.7 596 16 Q7W001
18 57 38.0 382 16 Q7UDT7
19 57 38.0 391 16 Q83MN5
20 55 36.7 479 16 Q25715
21 55 36.7 479 16 Q9ZM80
22 54 36.0 533 9 Q94MX6
23 53 35.3 591 16 Q8EP50
24 53 35.3 682 10 Q9M3E5
25 52 34.7 676 10 Q8H411
26 52 34.7 858 5 Q17647
27 52 34.7 860 5 Q95NM4
28 51 34.0 283 16 Q8ZMG6
29 51 34.0 271 16 O50860
30 51 34.0 698 10 Q7XWV8
31 50 33.3 389 10 Q8GZY3
32 50 33.3 433 10 Q84SN3
33 50 33.3 598 16 Q88VQ8
34 50 33.3 599 16 Q8EP55
35 50 33.3 647 10 Q8H413
36 50 33.3 701 10 Q852A9
37 49.5 33.0 735 10 Q9SW88
38 49.5 33.0 168 10 Q22625
39 49.5 33.0 848 16 Q82WX2
40 49.5 33.0 2021 5 O62191
41 49 32.7 197 12 P89214
42 49 32.7 382 16 Q7U6Q2
43 49 32.7 561 5 Q8IJ58
44 49 32.7 587 10 Q9LYU1
45 49 32.7 613 10 Q9M1G4

```

ALIGNMENTS

RESULT 1

```

015762 PRELIMINARY; PRT; 64 AA.
ID O15762
AC O15762;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP-3.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=94348085; PubMed=8068948;
RA Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Aikawa M., Filgueira M.-C., Tartar A., Drulhe P.;
RT "Merozoite surface protein-3: a malaria protein inducing antibodies
RT that promote Plasmodium falciparum killing by cooperation with blood
RT monocytes.";
RL Blood 84:1594-1602(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RA Oeuvray C., Bouharoun-Tayoun H., Gras-Masse H., Bottius E., Kaidoh T.,
RA Aikawa M., Filgueira M.-C., Tartar A., Drulhe P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF024624; AAB81610.1; -.
FT NON_TER 1
FT NON_TER 64
SQ SEQUENCE 64 AA; 7326 MW; 02E0ABF614C15A3B CRC64;

```

Query Match 100.0%; Score 150; DB 5; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKASDYVILGWFGGVPEHKKEEN 27

```
Db 18 AKEASSYDYLWGFEFGGVPEHKKEEN 44
RESULT 2
Q9NFV9 ID Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252287; CAB85901.1; -.
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; 36916 MW; C5B045DB5E21A159 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEFGGVPEHKKEEN 27
Db 174 AKEASSYDYLWGFEFGGVPEHKKEEN 200
RESULT 3
Q25995 ID Q25995 PRELIMINARY; PRT; 354 AA.
AC Q25995;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Antigen.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favalaro J.M.,
RA Thompson J.K., Marshall V.M., Coppell R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with Plasmodium falciparum merozoites.";
RL Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NF54;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; L28825; AAC09377.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEFGGVPEHKKEEN 27
Db 184 AKEASSYDYLWGFEFGGVPEHKKEEN 210
RESULT 5
Q95P15 ID Q95P15 PRELIMINARY; PRT; 361 AA.
AC Q95P15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVO;
RA Hisaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
RT "Merozoite Surface Protein 3 and Protection Against Malaria in Actus RT Monkeys.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY044180; AAK94780.1; -.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 361;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKEASSYDYLWGFEFGGVPEHKKEEN 27
```



```
Db 190 AKEASSYDYLIGWFGGGVPEHKKEEN 216

RESULT 6
Q25705 ID Q25705 PRELIMINARY; PRT; 379 AA.
AC Q25705
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
FM Mol. Biochem. Parasitol. 90:21-31(1997).
L EMBL; U08851; AAC47831.1; -.
SQ SEQUENCE 379 AA; 43344 MW; DC7AF106887C8AA0 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 210 AKEASSYDYLIGWFGGGVPEHKKEEN 236

RESULT 7
Q25706 ID Q25706 PRELIMINARY; PRT; 379 AA.
AC Q25706
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RC STRAIN=cs12;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
FM Mol. Biochem. Parasitol. 90:21-31(1997).
L EMBL; U08852; AAC47832.1; -.
SQ SEQUENCE 379 AA; 43302 MW; ABF9D54E1ED91A24 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 210 AKEASSYDYLIGWFGGGVPEHKKEEN 236

RESULT 8
Q9U6C4 ID Q9U6C4 PRELIMINARY; PRT; 379 AA.
AC Q9U6C4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polymorphic antigen.
```

```
GN MSP-3.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RC STRAIN=FCC1/HN;
RA Li X.R., Yu X.B., Luo S.H., Shan Z.X., Fang J.M., Xu J.;
RT "Sequence of Plasmodium falciparum secreted polymorphic antigen
   gene.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF188190; AAF04099.1; -.
SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 210 AKEASSYDYLIGWFGGGVPEHKKEEN 236

RESULT 9
Q26019 ID Q26019 PRELIMINARY; PRT; 380 AA.
AC Q26019
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
   Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
   Plasmodium falciparum merozoites.";
FM Mol. Biochem. Parasitol. 68:53-67(1994).
L EMBL; L07944; AAC09378.1; -.
SQ SEQUENCE FROM N.A.
RC STRAIN=FC27;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RT "Conservation of structural motifs and antigenic diversity in the
   Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
FM Mol. Biochem. Parasitol. 90:21-31(1997).
L EMBL; L07944; AAC09378.1; -.
DR PDB; 1PSW; 07-FEB-95.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 380 POLYMORPHIC ANTIGEN.
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 150; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AKEASSYDYLIGWFGGGVPEHKKEEN 27
Db 211 AKEASSYDYLIGWFGGGVPEHKKEEN 237

RESULT 10
Q9U0G0 ID Q9U0G0 PRELIMINARY; PRT; 346 AA.
AC Q9U0G0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Polymorphic antigen.
```

```

DN 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20416497; PubMed=10960178;
RA Okeno D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252286; CAB65754.1; -.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;

Query Match 96.7%; Score 145; DB 5; Length 346;
Best Local Similarity 96.3%; Pred. No. 6.4e-13;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AKESYDYILGWFGGVPHEKKEEN 27
Db |||||:|||||:|||||:|||||:|||||
187 AKESYNYILGWFGGVPHEKKEEN 213

RESULT 11
Q9GS22 PRELIMINARY; PRT; 371 AA.
AC Q9GS22;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Merozoite surface protein 6.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21106479; PubMed=11166390;
RA Trucco C., Fernandez-Reyes D., Howell S., Stafford W.H.,
RA Scott-Finnigan T.J., Grainger M., Ogun S.A., Taylor W.R., Holder A.A.;
RT "The merozoite surface protein 6 gene codes for a 36 kDa protein
RT associated with the Plasmodium falciparum merozoite surface protein-1
RT complex.";
RL Mol. Biochem. Parasitol. 112:91-101(2001).
DR EMBL; AY007721; AAG15394.1; -.
DR InterPro; IPR001680; WD40.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
SQ SEQUENCE 371 AA; 42275 MW; 4788BC1FAF965CD6 CRC64;

Query Match 44.3%; Score 66.5; DB 5; Length 371;
Best Local Similarity 63.6%; Pred. No. 0.23;
Matches 14; Conservative 0; Mismatches 3; Indels 5; Gaps 1;

QY 10 ILGWFGGVP-----EHKKEE 26
Db |||||:|||||:|||||:|||||:|||||
182 ILGWFGGAPQNGAAEDKKE 203

RESULT 12
Q8IJ54 PRELIMINARY; PRT; 371 AA.
AC Q8IJ54;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Merozoite surface protein 6.
GN PF10_0346.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;

QY 10 ILGWFGGVP-----EHKKEE 26
Db |||||:|||||:|||||:|||||:|||||
182 ILGWFGGAPQNGAAEDKKE 203

RESULT 13
Q8IJ48 PRELIMINARY; PRT; 405 AA.
AC Q8IJ48;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PF10_0352.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=2255705; PubMed=12369864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eissen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511(2002).
DR EMBL; AB014834; AAN35549.1; -.
KW Hypothetical protein.
SQ SEQUENCE 405 AA; 46729 MW; 5C6BC2106B94F5FD CRC64;

Query Match 44.0%; Score 66; DB 5; Length 405;
Best Local Similarity 47.6%; Pred. No. 0.3;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 YDYILGWFGGVPHEKKEEN 27
Db |:::|||||:|||||:|||||
232 YNHFAWEIGGAPTYKPKENN 252

RESULT 14
O69947

```

ID 069947 PRELIMINARY; PRT; 1238 AA.
 AC 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein SC06572.
 GN SC06572 OR SC3F9.07.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX Seeger K.J., Harris D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBSJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBSJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Kienast H., Hopwood D.A.;
 RL "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RL "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 Nature 417:141-147(2002).
 DR EMBL: AL939128; CAA19630.1; -;
 DR F01: T34929; T34929.
 DR GO: GO:0030246; F:carbohydrate binding; IEA.
 DR InterPro: IPR006584; CBD IV.
 DR InterPro: IPR005084; CBM 6.
 DR InterPro: IPR008979; Gal_bind_like.
 DR InterPro: IPR00601; PKD.
 DR Pfam: PF03422; CBM_6; 1.
 DR Pfam: PF00801; PKD; 1.
 DR SMART: SM00606; CBD IV; 1.
 DR SMART: SM00089; PKD; 1.
 DR PROSITE: PS00093; PKD; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 1238 AA; 133463 MW; ADAIE2F9D56C9410 CRC64;
 Query Match 40.7%; Score 61; DB 16; Length 1238;
 Best Local Similarity 44.8%; Pred. No. 5.7;
 Matches 13; Conservative 2; Mismatches 8; Indels 6; Gaps 1;
 QY 1 AKRASSYDYILGWFGG-----VPEHK 23
 DB 531 AKDADSPDLYTGWDFGDTGKGLTPTHK 559
 RESULT 15
 Q7WR42
 ID Q7WR42 PRELIMINARY; PRT; 596 AA.

AC Q7WR42;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Aspartyl-tRNA synthetase (EC 6.1.1.12).
 GN ASPS OR TLS OR BB0115.
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271;
 RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagsels K.,
 RA Leather S., Moule S., Norbertzak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RL "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica.";
 Nat. Genet. 35:32-40(2003).
 DR EMBL: BX640437; CAE30616.1; -;
 KW Aminocyl-tRNA synthetase; Ligase; Complete proteome.
 SQ SEQUENCE 596 AA; 67165 MW; F74DE904F8DDA448 CRC84;
 Query Match 38.7%; Score 58; DB 16; Length 596;
 Best Local Similarity 54.2%; Pred. No. 7;
 Matches 13; Conservative 4; Mismatches 5; Indels 2; Gaps 2;
 QY 4 ASSYDYL-GWFGG-VPEHKX 25
 DB 478 AKAYDMVLNGWIEGGGVRIHREE 501
 Search completed: September 22, 2004, 18:09:07
 Job time : 30.562 secs

W. E. B. DUBOIS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 22, 2004, 18:01:47 ; Search time 29.6198 Seconds
(without alignments)
298.263 Million cell updates/sec

Title: US-10-774-602-13
Perfect score: 147
Sequence: 1 PEHKKENMLSHLYVSSKOKENISKEND 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

1 number of hits satisfying chosen parameters: 1017041

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	100.0	329	5 Q9NFV9	Q9NFV9 plasmodium
2	147	100.0	346	5 Q9UG00	Q9UG00 plasmodium
3	147	100.0	354	5 Q25995	Q25995 plasmodium
4	147	100.0	354	5 Q81J55	Q81J55 plasmodium
5	147	100.0	361	5 Q95P15	Q95P15 plasmodium
6	147	100.0	379	5 Q25705	Q25705 plasmodium
7	147	100.0	379	5 Q25706	Q25706 plasmodium
8	147	100.0	379	5 Q9UGC4	Q9UGC4 plasmodium
9	147	100.0	380	5 Q26019	Q26019 plasmodium
10	143	97.3	64	5 Q15762	Q15762 plasmodium
11	70.5	48.0	405	5 Q81J48	Q81J48 plasmodium
12	64	43.5	602	5 Q81B95	Q81B95 plasmodium
13	62	42.2	424	5 Q81J53	Q81J53 plasmodium
14	58	39.5	7170	5 Q81L30	Q81L30 plasmodium
15	57	38.8	2249	5 Q81IY8	Q81IY8 plasmodium
16	55	37.4	425	6 Q95221	Q95221 oryctolagus

17	55	37.4	448	11	Q80V72	mus musculus
18	55	37.4	524	17	Q8TR11	mus musculus
19	55	37.4	527	11	Q8RON3	mus musculus
20	55	37.4	551	11	Q811K1	mus musculus
21	55	37.4	1056	11	Q8COP7	mus musculus
22	54.5	37.1	116	12	Q80IM0	eupatorium
23	54.5	37.1	584	5	Q8IDV8	plasmodium
24	54.5	37.1	1155	10	Q9FX38	arabidopsis
25	54.5	37.1	3063	5	Q81284	plasmodium
26	54	36.7	214	10	Q9AX37	oryza sativ
27	54	36.7	371	5	Q9GS22	plasmodium
28	54	36.7	371	5	Q81J54	plasmodium
29	54	36.7	839	9	Q64076	bacterioph
30	54	36.7	839	16	Q31945	bacillus su
31	54	36.7	983	5	Q8S9J2	encephalito
32	53	36.1	92	10	Q9LJ31	oryza sativ
33	53	36.1	351	4	Q9UFU6	homo sapien
34	53	36.1	426	4	Q9UHH9	homo sapien
35	53	36.1	426	4	Q9H4P7	homo sapien
36	53	36.1	2418	5	Q811R9	plasmodium
37	53	36.1	2625	5	Q8MMZ9	dictyosteli
38	53	36.1	3242	5	Q9V4N8	drosophila
39	52.5	35.7	600	5	Q77355	plasmodium
40	52.5	35.7	1079	10	Q93XF7	zea mays (m
41	52.5	35.7	1132	5	Q815C3	plasmodium
42	52.5	35.7	1435	5	Q81LL4	plasmodium
43	52.5	35.7	1883	5	Q81LI0	plasmodium
44	52	35.4	63	16	Q81RM5	bacillus an
45	52	35.4	315	5	Q97304	plasmodium

ALIGNMENTS

RESULT 1

Q9NFV9 ID Q9NFV9 PRELIMINARY; PRT; 329 AA.
AC Q9NFV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium falciparum (isolate 7G8).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=57266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=7G8;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
Plasmodium reichenowi and Plasmodium falciparum."
RL Mol. Biochem. Parasitol. 109:185-188(2000).
DR EMBL; AJ252287; CAB85901.1; -;
FT NON_TER 1
FT NON_TER 329
SQ SEQUENCE 329 AA; 36916 MW; C5B045DB5E21A159 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 329;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKENMLSHLYVSSKOKENISKEND 28
|
Db 193 PEHKKENMLSHLYVSSKOKENISKEND 220
|

RESULT 2

Q9UG00 ID Q9UG00 PRELIMINARY; PRT; 346 AA.
AC Q9UG00;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
GN MSP3.
OS Plasmodium reichenowi.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5854;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;
RX MEDLINE=20416497; PubMed=10960178;
RA Okenu D.M.N., Thomas A.W., Conway D.J.;
RT "Allelic lineages of the merozoite surface protein 3 gene in
RT Plasmodium reichenowi and Plasmodium falciparum.";
RL Mol. Biochem. Parasitol. 109:185-188 (2000).
DR EMBL, AJ252286; CAB65754.1; -.
FT NON_TER 1
FT NON_TER 346
SQ SEQUENCE 346 AA; 39127 MW; A804B96BDFAF010 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 346;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 206 PEHKKEENMLSHLYVSSKDKENISKEND 233

RESULT 3
Q25995 ID Q25995 PRELIMINARY; PRT; 354 AA.
AC Q25995;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Antigen.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5843;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;
RX STRAIN=NF54;
RC MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
RA Thompson J.K., Marshall V.M., Coppe R.L., Kemp D.J., Anders R.F.;
RT "Molecular variation in a novel polymorphic antigen associated with
RT Plasmodium falciparum merozoites.";
RL Mol. Biochem. Parasitol. 90:21-31 (1997).
DR EMBL, L28825; AAC09377.1; -.
SQ SEQUENCE 354 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 203 PEHKKEENMLSHLYVSSKDKENISKEND 230

RESULT 4
Q81J55 ID Q81J55 PRELIMINARY; PRT; 354 AA.
AC Q81J55;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Merozoite surface protein 3.
GN PF10_0345.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;
RX STRAIN=3D7;
RC MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum.";
RL Nature 419:498-511 (2002).
DR EMBL, AE014834; AAN35542.1; -.
SQ SEQUENCE 361 AA; 40119 MW; 3A7256152F48B527 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 354;
Best Local Similarity 100.0%; Pred. No. 2.1e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 203 PEHKKEENMLSHLYVSSKDKENISKEND 230

RESULT 5
Q95P15 ID Q95P15 PRELIMINARY; PRT; 361 AA.
AC Q95P15;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]_SEQUENCE FROM N.A.
RP SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;
RX STRAIN=FVO;
RA Hissaeda H., Saul A., Long C.A., Miller L.H., Stowers A.W.;
RT "Merozoite Surface Protein 3 and Protection Against Malaria in Aotus
RT Monkeys.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AY044180; AAK94780.1; -.
FT NON_TER 1
FT NON_TER 361
SQ SEQUENCE 361 AA; 41163 MW; 6127A3041587BA74 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 361;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 209 PEHKKEENMLSHLYVSSKDKENISKEND 236

RESULT 6
Q25705 ID Q25705 PRELIMINARY; PRT; 379 AA.
AC Q25705;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

```

RT gene. ";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RD EMBL; AF18190; AA04099.1; -.
SQ SEQUENCE 379 AA; 43316 MW; C152A54E1F9D5F25 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.3e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 229 PEHKKEENMLSHLYVSSKDKENISKEND 256
|||||
|||||

RESULT 9
Q26019 PRELIMINARY; PRT; 380 AA.
ID AC Q26019
AC Q26019;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polymorphic antigen precursor.
DE Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCB_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=FC27;
RX MEDLINE=95198774; PubMed=7891748;
RA McColl D.J., Silva A., Foley M., Kun J.F., Favaloro J.M.,
RA Thompson J.K., Marshall V.M., Coppel R.L., Kemp D.J., Anders R.F.;
RA "Molecular variation in a novel polymorphic antigen associated with
RA Plasmodium falciparum merozoites.";
RT Mol. Biochem. Parasitol. 68:53-67(1994).
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=FC27;
RX MEDLINE=98156743; PubMed=9497029;
RA McColl D.J., Anders R.F.;
RA "Conservation of structural motifs and antigenic diversity in the
RT Plasmodium falciparum merozoite surface protein-3 (MSP-3).";
RL Mol. Biochem. Parasitol. 90:21-31(1997).
DR EMBL; I07944; AAC09378.1; -.
DR PDB; 1PSM; 07-FEB-95.
KW SIGNAL.
FT SIGNAL.
FT CHAIN
FT SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;
SQ SEQUENCE 380 AA; 43290 MW; 0986CA1393094CA2 CRC64;

Query Match 100.0%; Score 147; DB 5; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.3e-11;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEHKKEENMLSHLYVSSKDKENISKEND 28
Db 230 PEHKKEENMLSHLYVSSKDKENISKEND 257
|||||
|||||

RESULT 10
Q15762 PRELIMINARY; PRT; 64 AA.
ID AC Q15762
AC Q15762;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Merozoite surface protein 3 (Fragment).
DE MSP-3.
OS Plasmodium falciparum (isolate NF54).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCB_TaxID=5843;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=NF54;
RC

```

[illegible]


```

RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sulston J.E., Craig A., Newbold C., Barrell B.G;
RT "Sequence of Plasmodium falciparum Chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531 (2002) .
RL EMBL; AL034557; CAD49133.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2249 AA; 271088 MW; 6EF04691C22FB787 CRC64;

Query Match 38.8%; Score 57; DB 5; Length 2249;
Best Local Similarity 47.8%; Pred. No. 99;
Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 HKKEENMLSHLYVSSKDKENISK 25
    ||| ||| | | | | |
Db 882 HKIESNNLTHINQKEDKKNISK 904

```

100